STRATEGIC EXPERIMENTATION WITH ERLANG BANDITS

SVETLANA BOYARCHENKO*

* Department of Economics, The University of Texas at Austin,
2225 Speedway Stop C3100, Austin, TX 78712, U.S.A.
e-mail: sboyarch@eco.utexas.edu

Abstract. Risks related to events that arrive randomly play important role in many real life decisions, and models of learning and experimentation based on two-armed Poisson bandits addressed several important aspects related to strategic and motivational learning in cases when events arrive at jump times of the standard Poisson process. At the same time, these models remain mostly abstract theoretical models with few direct economic applications. We suggest a new class of models of strategic experimentation which are almost as tractable as exponential models, but incorporate such realistic features as dependence of the expected rate of news arrival on the time elapsed since the start of an experiment and judgement about the quality of a "risky" arm based on evidence of a series of trials as opposed to a single evidence of success or failure as in exponential models with conclusive experiments. We demonstrate that, unlike in the exponential models, players may stop experimentation before the first failure happens. Moreover, *ceteris paribus*, experimentation in a model with breakthroughs may last longer than experimentation in the corresponding model with failures.

Keywords: stopping time games, Erlang distribution, strategic experimentation

JEL: C73, C61, D81

1. Introduction

This paper suggests a new class of learning and experimentation models based on *Erlang bandits*. We show that such models are more realistic than models based on Poisson bandits and generate qualitatively new results.

In many situations in real life, it is necessary to quantify risks related to events that arrive at random times, as well as frequency of their arrivals. For example, in finance, it is necessary to evaluate default risks of borrowers or assets and default rates. In pharmaceuticals, it is necessary to evaluate possible side effects of a new

I am thankful for discussions to participants of research seminars at The Center for Rationality in the Hebrew University of Jerusalem, at the University of Tel Aviv, at Haifa University, at the Research University Higher School of Economics, Moscow and St. Petersburg. I am especially grateful to Ehud Lehrer and Eilon Solan for pointing out mistakes in the previous version of the paper and making useful suggestions. I am also thankful for discussions to Max Stinchcombe. The usual disclaimer applies.

drug or its efficiency. In any kind of sponsored research, sponsors have to figure out the probability of success, and so on. At the same time, it is also important to take into consideration existence of strategic partners (or adversaries) - what sort of information one could learn from them or which pieces of information one would disclose.

An extensive literature on learning and experimentation based on so called Poisson bandits addresses these sorts of issues. A standard "two-armed" bandit is an attempt to describe a hypothetical experiment in which a player faces two slot machines; the quality of one of the slots is known ("safe" arm), and the other one ("risky" arm) may be "good" or "bad." In case of so called "conclusive" experiments - the first event observed on the "risky" arm reveals its quality completely, so the experiment is over, when the first ("conclusive") success or failure is observed. Bandit models were successfully used in various settings in economics, for example, learning and matching in labor markets, monopolist pricing with unknown demand, choice between R&D projects, or financing of innovations (see, e.g., [1, 2, 3, 4, 5, 6, 19, 27, 29, 30] and references therein).

Models of strategic learning and experimentation extend "two-armed" bandit experiments to a setting where several players face copies of the same slot machine. Players then learn about the quality of the risky arm not only from outcomes of their own experiments, but also from their colleagues. In models of strategic experimentation, it is common to assume away payoff externalities and focus on information externalities, the role of information, and, in more advanced settings, on design of information. See, for example, [7, 10, 17, 18, 20, 21]. Recent developments include (but are not limited to) correlated risky arms as in Klein and Rady [22] and Rosenberg et al. [28], or private payoffs as in Heidhues et al. [14] and Rosenberg et al. [28], or departures from Markovian strategies as in Hörner et al. [17]. For other developments and an excellent comprehensive review of the literature see Hörner and Skrzypacz [15] and references therein.

In continuous time models, the payoff generated by the "risky" arm of a "two-armed" bandit follows a certain continuous time stochastic process whose parameters are not known. For example, Bolton and Harris [7] model the unknown payoff as a Brownian motion with unknown drift and known variance in a model of strategic experimentation. Decamps et al. [11, 12] study timing a fixed size investment into a risky project with the payoff generated by a Brownian motion with unknown drift and known variance. Keller et al. [18], Keller and Rady [20, 21] use a Poisson process with unknown rate of arrival to model the risky arm. Decamps and Mariotti [10] study a duopoly model of investment where a signal about the quality of the project is modeled as a Poisson process. Cohen and Solan [8] bridge the gap between the Brownian motion and Poisson bandits and consider "two-armed" bandits, where the risky arm yields stochastic payoffs generated by a Lévy process.

Poisson bandits models with random costly breakdowns ("bad" news models) differ significantly from similar models with random profitable breakthroughs ("good" news models). The main distinction in these models arises from the fact that updating of beliefs when no news arrive moves in the opposite directions in case of breakdowns and breakthroughs. If nothing happens in the model with potentially "bad" news, players become more and more optimistic about the quality of the "risky" arms, and the expected rate of arrival of "bad" news decreases over time. Therefore, if it is optimal to start experimentation at the prior beliefs levels, experimentation never stops until the first breakdown occurs (if ever). On the contrary, if no successes arrive in the model with potentially "good" news, players become more and more pessimistic about the quality of the risky arms, and the expected rate of arrival of "good" news decreases over time, therefore, experimentation always stops in finite time, unless the first success arrives earlier. Note that experimentation levels in "good" news models are low.

The main feature of exponential bandits models is that the rate of arrival of "good" or "bad" news over a time interval $(t, t + \Delta t)$ is independent of t. This is convenient for tractability, but not very realistic. Possibly, this unrealistic feature is one of the reasons why Poisson bandits experimentation models remain mostly abstract theoretical models with few direct economic applications.

Indeed, even if a lender assigns a high prior belief to potential default of a borrower, she would hardly expect the borrower being equally likely to default immediately upon the loan initiation or some time after. Moreover, defaults typically do not happen completely "out of the blue" - they are preceded by several smaller events such as missed monthly loan payments, for example.

The 2010 Deepwater Horizon disaster, which seemed to be quite unexpected to unsophisticated observers, was preceded by early signs of degradations in the safety systems of BP, such as a series of explosions and fires at the Grangemouth facility in Scotland between 1987-2001.

When a researcher starts a new research project, sponsors hardly expect her to generate immediate success even when they have high beliefs in the quality of the project. At the same time, occurrence of a single success may be inconclusive - it has to be replicated in several similar trials for a project to be labeled as a "success." For example, recently, drug company Merck had to halt the late-stage trial of its promising Alzheimer's drug verubecestat, after an independent study demonstrated that the drug was not working as expected.

We propose a class of experimentation models which are almost as tractable as exponential models and can reflect such realistic features that rates of news arrival may depend on the amount of time elapsed since an experiment had started and that better assessment of the quality of the tested "arm" may require several consecutive random events as opposed to a single random success or failure. Namely, we consider scenario, where two agents experiment with a project of unknown quality (we consider separately "bad" and "good" news models) and assume that random times of news arrival are Erlang-2 random variables. In this case, the expected rate of arrival of the news is a non-monotone function of time - it grows from zero (at the moment when experiment starts) to a certain maximal level, then starts decreasing and behaving more and more as a negative exponential as the experiment "grows older." Thus, for some time after the start of the experiment, the beliefs about the quality of the "risky" arm and the expected rate of arrival of "bad" or "good" news move in the opposite directions. In particular, this may cause the players stop experimentation in finite time even in the model with breakdowns. The longer the experimenter observes no "bad" news, the more optimistic she becomes about the quality of the risky arm, but at the same time, the anticipated rate of arrival of a failure also grows, and these two opposite movements makes stopping before the first breakdown optimal. If such stopping is optimal, it happens where the expected rate of arrival is increasing. As opposed to this, in a model with breakthroughs, experimentation does not stop while the expected rate of arrival of "good" news keeps growing, even though experimenters become more pessimistic about the quality of the "risky" arm. Khan and Stinchcombe [23] find similar results in semi-Markovian decision theory. Namely, they identify two classes of situations in which delay in decision systems is optimal: in the first class delay is optimal when the hazard rate of further changes is increasing, and in the second class, delay is optimal when the hazard rate is decreasing.

Stopping regions in exponential bandits are described in terms of cutoff beliefs. In many situations, "internal" experimentation happens: that is before complete switch from experimentation with the "risky" arm only to playing the "safe" arm only, players start diversifying their time between these two activities. Regions of "internal" experimentation (or "partial exit" from experimentation) are also described in terms of cutoff beliefs. We argue that a more relevant characterization of stopping rules is in terms of cutoff expected rates of the news arrivals. While in exponential models characterizations in terms of cutoff beliefs and cutoff expected rates of the news arrival are equivalent, they are not so in Erlang models. Furthermore, stopping rules based on expected rates of the news arrivals have clear economic interpretation. In a model with failures, a player stops experimentation when the marginal benefit from staying active equals the expected marginal cost. If the marginal benefit is higher than the expected marginal cost at any time, the player never stops before the first failure is observed (if ever). In a model with breakthroughs, a player stops experimentation when the expected marginal benefit from staying active equals the marginal cost. If parameters of a "good" and "bad" news models are such that the cutoff expected rates of the news arrival are the same, experimentation in the "good" news model lasts longer than in the "bad" news model, provided no news arrive in either model.

Another rather unrealistic feature of traditional Poisson bandit models is that the value of an outside option ("safe" arm payoff) is the same no matter whether this option is taken before any news arrive or after that. For example, one gets a higher value for a used car if there were no recalls or negative consumers' reports on this make. Similarly, if one abandons a research project that generated no positive outcome, the recovery value is smaller than in case of a successful project that ends, for example, in patenting a new invention. To make our model more realistic, we assume that if one or both players exit before any failure happens, the recovery value (the value of the outside option) they get is higher than the one which they can get if they exit after a failure had revealed the "bad" quality of the tested project. This loss in the recovery value may be due to a reputation loss or the cost of liquidation of the research facility after a costly accident, or the costs a pharmaceutical company has to pay in case patients testing new drugs develop side effects. While each of the two active players is equally likely to suffer a cost of failure, both of them suffer the loss of recovery value in case of a failure, no matter who incurs a failure. We show that when the loss in the recovery value is sufficiently high, the value of a single player is higher than a value of a player who exits first (the leader), but the value of a single player is lower than the value of the player who exits second (the follower). Similarly, in the "good" news model, we distinguish payoffs in case experimentation stops before the first success was observed and after the first success was observed. We also assume that the player who is the first one to observe a breakthrough gets higher payoff than the other player.

Depending on parameters of the model with failures, the following subgame perfect equilibria are possible in the corresponding stopping time game: (1) none of the players find it optimal to stop unless the first failure happens; (2) the leader stops in finite time, and the follower either exits later in finite time, or never unless the first failure happens; (3) a symmetric equilibrium, where the players do not stop until the optimal stopping time of the leader and then randomize in the interval between the leader's and the follower's optimal stopping times. Type (3) equilibrium is Pareto dominated by asymmetric equilibria of type (2), because in the former equilibrium both players get the leader's payoff, which is lower than the follower's payoff. If there is no loss in the recovery value in case of a failure, the follower optimal stopping time becomes the same as the stopping time of the leader, so type (2) and (3) equilibria collapse into one equilibrium, where the players stop simultaneously at the leader's optimal stopping time. Similar results hold in the model with breakthroughs.

The papers which are mostly close to our paper are Keller and Rady [20, 21] and Rosenberg et al. [28]. Keller and Rady [21] study the case of costly breakdowns that arrive at the jump times of Poisson processes which are independent. Rosenberg et al. [28] consider an irreversible exit problem in a model with breakthroughs with correlated risky arms both in the case when payoffs are public and private.

The rest of the paper is organized as follows. Section 2 considers a stopping time game in a "bad" news model. The detailed analysis of the game and construction of subgame perfect equilibria are in Section 3. A stopping game in a "good news" model is analyzed in Section 5. Generalizations for more general intensities of the news arrival (*humped bandits*) are outlined in Section 6. Section 7 concludes. Technical proofs are relegated to the appendix.

2. Conslusive failures

2.1. **The setup.** We consider the game of timing, characterized by the following structure. Time $t \in \mathbb{R}_+$ is continuous, and the discount rate is $r > 0$. Two symmetric players experiment with risky projects, such as a nuclear technology, a defaultable loan, or a new drug. The quality of the project depends on the state of nature $\theta \in \{0, 1\}$. If $\theta = 0$, the project is "good," which means that it never fails. If $\theta = 1$, the project is "bad," which means that the player experimenting with this project may incur costs when the project fails, which may happen at random times. Assume that the (lump-sum) cost in case of a failure is $\hat{C} > 0$. The quality of the project is not known initially. We leave for the future study the case when the players can have projects of different types, and the types may be positively of negatively correlated as in Rosenberg et al. [28] or Klein and Rady [22], and assume that the quality of the project is the same for both players, so that if a failure happens to one of the players' projects, both players know that their projects are "bad." The initial common prior assigns probability $\bar{\pi} \in (0, 1)$ to $\theta = 0$.

An active experimenting player gets a constant revenue stream $rR > 0$ as long as no failure had been observed. This stream can be viewed, for example, as sponsored research contributions, or revenue generated by a project net of insurance costs, or mortgage payments. For simplicity, we assume that if the project is "bad", then after the first failure, the stream of revenues disappears (e.g., the sponsor withdraws support from a pharmaceutical company as soon as a side effect of a new drug is observed; the insurance company increases the premium to the extent that offsets the revenue stream of a faulty technology; a borrower is not able to make monthly payments after a default, etc). Given this assumption, experimentation after the first failure becomes non-profitable, so the players stop experimenting as soon as they learn that the quality of the project is "bad."

Let τ_i denote the random time of the first failure of player i's project if $\theta = 1$. We assume that τ_i and τ_j are i.i.d. $Erlang(2, \lambda)$ random variables[1]. Note that the expected time until the first failure is $2/\lambda$.

W.l.o.g. assume that the game starts at $t = 0$. At each point $t \geq 0$, player $i \in \{1, 2\}$ may make an irreversible stopping decision conditioned on the history of the game. The value of an outside option (recovery value) in case of exit is $S \in [0, R)$ if exit happens before the first failure is observed, and $S - L$ if it happens at the time of the first failure or later. Here $L \leq S$ represents the loss in the recovery value due to the fact that the "bad" quality of the project has been revealed. If $S = 0$, then $L \leq 0$ can be viewed as an additional liquidation cost of the research facility. Set $C = \hat{C} + L$.

In the current setting, we consider the case when all payoffs, parameters of the Erlang distribution, and the players' actions are public information. W.l.o.g. assume

[1] The p.d.f. of $Erlang(k, \lambda)$ distribution for $k \geq 1$ and $\lambda > 0$ is given by $f(t) = \lambda^k t^{k-1} e^{-\lambda t}/(k-1)!$.

that the game starts at $t = 0$. At each point $t \geq 0$, player $i \in \{1, 2\}$ may make an irreversible stopping decision conditioned on the history of the game. At any $t \geq 0$, the history of the game includes observations of all failures (including the empty set if no failures were observed by the players up to time t) and the actions of the players. As far as the actions are concerned, only two sorts of histories matter in the stopping game: (i) both players are still in the game; (ii) at least one player exited the game.

Let $T_i \in \mathbb{R}_+$ denote the exit time of player i. Define the function

$$\tilde{t}_i(t) = \begin{cases} T_i, & \text{if } T_i \leq t, \\ \infty, & \text{otherwise.} \end{cases}$$

Let τ_i^s denote a random time, when a failure of player i's project occurred for the s^{th} time. The history of observations at any $t \geq 0$ is

$$O_t = \{(\tau_1^{s'})\}_{s' \leq t \wedge T_1} \cup \{(\tau_2^{s''})\}_{s'' \leq t \wedge T_2}.$$

If $O_t = \{\emptyset\}$, then a typical history at time t is $h_t(O_t, \tilde{t}_1(t), \tilde{t}_2(t))$. If $T_i < T_j$, we call player i the leader, and player j the follower. If $O_t \neq \{\emptyset\}$, the game is over.

Definition 2.1. A simple strategy for player $i \in \{1, 2\}$ in the game starting at $t = 0$ is a function

$$q_i : [0, +\infty) \to [0, 1],$$

which is non-increasing, left-continuous with right limits (LCRL); and $q_i(0) = 1$.

Following Laraki et al. [24], and Dutta and Rustichini [13], for any time $t > 0$, define a proper subgame as the timing game that starts at the end of the history h_t.

Definition 2.2. A simple strategy for player $i \in \{1, 2\}$ in a subgame game starting at $t > 0$ is a function

$$q_i^t : [t, +\infty) \to [0, 1],$$

which is non-increasing, left-continuous with right limits (LCRL); and $q_i^t(t) = 1$.

It follows from Definitions 2.1 and 2.2, that $q_i(t)$ is the probability that player i did not exit at time t or earlier, and $q_i^t(s)$ is the probability that player i did not exit during $[t, s]$, conditioned on being active at t. In either of the definitions above, we allow for the case $q_i(+\infty) > 0$ and $q_i^t(+\infty) > 0$, which means that player i decides not to exit ever with the positive probability $q_i(+\infty)$ and $q_i^t(+\infty)$, respectively, unless a failure happens. Thus, the probability that the agent will not exit is $q_i(+\infty)\bar{\pi}$ and $q_i^t(+\infty)\bar{\pi}$, respectively.

Definition 2.3. A simple strategy of player i is called consistent, if for any $0 \leq t \leq t' \leq s$,

$$q_i^t(s) = q_i^t(t') q_i^{t'}(s). \tag{2.1}$$

2.2. **Evolution of beliefs, "survival" probabilities, and rates of arrival.** Let $\tau \sim Erlang(2, \lambda)$. Use the following notation:

$$p_\lambda(t) = \mathrm{prob}(\tau > t) = \int_t^\infty \lambda^2 s e^{-\lambda s} ds = (\lambda t + 1)e^{-\lambda t}.$$

Let τ_i and τ_j be i.i.d. $Erlang(2, \lambda)$ random variables. Then

$$p_{2\lambda}(t) := \mathrm{prob}(\tau_i > t)\mathrm{prob}(\tau_j > t) = p_\lambda^2(t) = (\lambda t + 1)^2 e^{-2\lambda t}.$$

Consider any $t \geq 0$ s.t. $O_t = \{\emptyset\}$. Given the prior belief $\bar{\pi}$, let $p_d(\bar{\pi}, t)$ denote the "survival" probability (i.e., the probability of the event that no failure happens before time t) if two players experiment with the project of unknown quality. Then $p_d(\bar{\pi}, t)$ is given by

$$p_d(\bar{\pi}, t) = \bar{\pi} + (1 - \bar{\pi})p_{2\lambda}(t) = \bar{\pi} + (1 - \bar{\pi})(\lambda t + 1)^2 e^{-2\lambda t}. \tag{2.2}$$

Function

$$[0, \infty) \ni t \mapsto p_d(\bar{\pi}, t) \in \mathbb{R}_+$$

defines the time-inhomogeneous Poisson process with the rate of arrival

$$\lambda_d(\bar{\pi}, t) = -\frac{p'_d(\bar{\pi}, t)}{p_d(\bar{\pi}, t)} = \frac{2(1 - \bar{\pi})\lambda^2 t(1 + \lambda t)e^{-2\lambda t}}{\bar{\pi} + (1 - \bar{\pi})(\lambda t + 1)^2 e^{-2\lambda t}}. \tag{2.3}$$

Similarly, given the prior belief $\bar{\pi}$, let $p_s(\bar{\pi}, t)$ denote the "survival" probability if one player experiments with the project of unknown quality. Then $p_s(\bar{\pi}, t)$ is given by

$$p_s(\bar{\pi}, t) = \bar{\pi} + (1 - \bar{\pi})p_\lambda(t) = \bar{\pi} + (1 - \bar{\pi})(\lambda t + 1)e^{-\lambda t}. \tag{2.4}$$

Function

$$[0, \infty) \ni t \mapsto p_s(\bar{\pi}, t) \in \mathbb{R}_+$$

defines the time-inhomogeneous Poisson process with the rate of arrival

$$\lambda_s(\bar{\pi}, t) = -\frac{p'_s(\bar{\pi}, t)}{p_s(\bar{\pi}, t)} = \frac{(1 - \bar{\pi})\lambda^2 t e^{-\lambda t}}{\bar{\pi} + (1 - \bar{\pi})(\lambda t + 1)e^{-\lambda t}}. \tag{2.5}$$

Let $\pi_d(\bar{\pi}, t)$ (respectively, $\pi_s(\bar{\pi}, t)$) denote the belief that $\theta = 0$ at time t, conditioned on no failures happened before t if two (respectively, one) players experiment(s) with the project of unknown quality. Then the beliefs $\pi_d(\bar{\pi}, t)$ and $\pi_s(\bar{\pi}, t)$ are given by

$$\pi_d(\bar{\pi}, t) = \frac{\bar{\pi}}{p_d(\bar{\pi}, t)}, \tag{2.6}$$

$$\pi_s(\bar{\pi}, t) = \frac{\bar{\pi}}{p_s(\bar{\pi}, t)}. \tag{2.7}$$

Observe that, for any $t > 0$,

$$\pi_d(\bar{\pi}, t) = \pi_s\left(\pi_s(\bar{\pi}, t), t\right). \tag{2.8}$$

Remark 2.4. Equation (2.8) can be generalized for the case of n players experimenting with projects of unknown quality with i.i.d. random times of arrival of failures. Namely, let $\pi_k(\bar{\pi}, t)$ be the belief that $\theta = 0$ at time t, conditioned on no failures happened before t if k players experiment. Then, for any $t > 0$,

$$\pi_n(\bar{\pi}, t) = \pi_1\left(\pi_{n-1}(\bar{\pi}, t), t\right). \tag{2.9}$$

To see, why, let $\hat{\pi}$ be a solution to

$$\pi_n(\bar{\pi}, t) = \pi_1\left(\hat{\pi}, t\right). \tag{2.10}$$

We write (2.10) as

$$\frac{\bar{\pi}}{\bar{\pi} + (1-\bar{\pi})p_\lambda^n(t)} = \frac{\hat{\pi}}{\hat{\pi} + (1-\hat{\pi})p_\lambda(t)}.$$

Equivalently,

$$\bar{\pi}(1-\hat{\pi})p_\lambda(t) = \hat{\pi}(1-\bar{\pi})p_\lambda^n(t),$$

whence

$$\hat{\pi} = \frac{\bar{\pi}}{\bar{\pi} + (1-\bar{\pi})p_\lambda^{n-1}(t)} = \pi_{n-1}(\bar{\pi}, t),$$

and (2.9) follows.

2.3. **Value functions and equilibrium.** Consider the game that starts at $t = 0$. Rates of arrival given by (2.3) and (2.5) are continuous and equal to zero at $t = 0$. Therefore, none of the players has yet stopped at the start of the game. As the rate of arrival $\lambda_d(\bar{\pi}, t)$ increases, it may become optimal for one or both players to quit. We will prove that, depending on parameters of the model, either the players do not stop until the first failure happens, or the stopping rules in pure strategies are of the threshold type - the players quit when the corresponding rates of arrival reach a certain threshold from below; in addition, there may exist a time interval such that (q_1^t, q_2^t) continuously decrease from one to zero on this interval. Given a strategy profile (q_1, q_2), one may observe the following outcomes: (i) none of the players stops before the first failure; (ii) only one of the players stops before the first failure; (iii) players stop before the first failure simultaneously; (iv) players stop before the first failure sequentially. The first two outcomes are possible if the ratio $r(R-S)/C$ is sufficiently high.

In order to define value functions of the players in this game, we will use the following version of the definition of the Riemann-Stieltjes integral.

Definition 2.5. Assume that the following conditions hold

- (i) $q, \Psi : [0, +\infty) \to \mathbb{R}$ are bounded LCRL functions;
- (ii) q is of finite variation;
- (iii) the singular continuous component of the Lebesgue decomposition of q is trivial;
- (iv) I is an interval of one of the following forms: $(a, b), [a, b), (a, b], [a, b], [a, +\infty), (a, +\infty)$, where $0 \le a < b < +\infty$ or a union of non-intersecting intervals of this form.

Define

$$\int_I \Psi(s)dq(s) = \int_I \Psi(s)q'(s)ds + \sum_{t_j\in I:\Delta q(t_j)\neq 0} \Psi(t_j)\Delta q(t_j), \qquad (2.11)$$

where $\Delta q(t_j) := q(t_j + 0) - q(t_j)$.

Let $G_i(\bar\pi, t)$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t > 0$. Let $F_i(\bar\pi, t)$ denote the expected value of player i if player j stopped at time t, and player i did not. From now on, we will consider the simple strategies q_1, q_2, whose singular continuous components of the Lebesgue decompositions are trivial. Given the strategy profile (q_i, q_j), the value of player i in the game that starts at $t = 0$ is

$$\begin{aligned} V_i(\bar\pi; q_i, q_j) \;&=\; \int_0^\infty e^{-rt} p_d(\bar\pi, t) G_i(\bar\pi, t) q_i(t) q_j(t)dt \qquad (2.12)\\ &+ \int_{\{t\geq 0 \;\mid\; \Delta q_i(t)=0\}} e^{-rt} p_d(\bar\pi, t) F_i(\bar\pi, t) q_i(t)(-dq_j(t))\\ &+ \int_0^\infty e^{-rt} p_d(\bar\pi, t) S q_j(t)(-dq_i(t)). \end{aligned}$$

Later we will show that $G_i(\bar\pi, \cdot)$ and $F_i(\bar\pi, \cdot)$ are continuous and have finite limits as $t \to \infty$, hence, $V(\bar\pi; q_i, q_j)$ is well-defined and finite. Note that the second integral in (5.2) takes into account jumps in q_j only, and the last integral takes into account jumps in q_i only as well as simultaneous jumps in q_i and q_j.

Definition 2.6. A strategy profile $\hat q = (\hat q_i, \hat q_j)$ is a Nash equilibrium for the game starting at $t = 0$, if for every $(i, j) \in \{(1, 2), (2, 1)\}$

$$V_i(\bar\pi; \hat q_i, \hat q_j) = \sup_{q_i} V_i(\bar\pi; q_i, \hat q_j).$$

A profile of consistent strategies $\hat q^t = (\hat q^t_i, \hat q^t_j)$ is a subgame perfect Nash equilibrium (SPE) if for every $t \geq 0$, $\hat q^t$ is a Nash equilibrium in the subgame that starts at t (when payoffs are discounted to time t).

3. Main steps of solution

Once one of the players has quitted experimentation, the other player faces a non-strategic stopping problem, which can be easily solved. Thus, when considering subgame perfect equilibria, we will first examine subgames when one of the players has stopped, and then move to subgames where neither player has quitted as yet. To simplify the notation, we suppress the dependence of value functions on the other player's strategy. Since the players are symmetric, we also drop the subscripts identifying the players.

3.1. **Follower's problem.** Consider a subgame that starts after the history such that no observations arrived, and only one of the players has stopped. Suppose, this happened at time t. Then the remaining player (the follower) chooses a strategy q_f^t satisfying conditions of Definition 2.2, which solves the following problem:

$$\begin{aligned} F(\bar{\pi}, t) &= \sup_{q_f^t} \left[\int_t^{\infty} e^{-r(t'-t)} q_f^t(t') \frac{p_s(\pi_s(\bar{\pi}, t), t')}{p_d(\bar{\pi}, t)} \left(rR + \lambda_s(\pi_s(\bar{\pi}, t), t')(S - C) \right) dt' \right. \\ &\quad \left. + S \int_t^{\infty} e^{-r(t'-t)} \frac{p_s(\pi_s(\bar{\pi}, t), t')}{p_d(\bar{\pi}, t')} (-dq_f^t(t')) \right], \end{aligned} \tag{3.1}$$

where the first integral is the expected present value (EPV) of the payoff while the follower is active, and the second integral is the EPV of the payoff, when the follower exits prior to the first failure.

Introduce the notation

$$A = \frac{r(R-S)}{C}, \tag{3.2}$$

$$\hat{A}_s(\pi_s(\bar{\pi}, t), t) = \max_{t' \geq t} \lambda_s(\pi_s(\bar{\pi}, t), t'), \tag{3.3}$$

$$\Phi(A, \bar{\pi}, t; T) = C \int_t^T e^{-rt'} p_s(\pi_s(\bar{\pi}, t), t') \left(A - \lambda_s(\pi_s(\bar{\pi}, t), t') \right) dt'. \tag{3.4}$$

Lemma 3.1. *The value of the follower, given by equation* (3.1)*, can be equivalently written as*

$$\begin{aligned} F(\bar{\pi}, t) &= S \\ &\quad + \frac{Ce^{rt}}{p_d(\bar{\pi}, t)} \sup_{q_f^t} \int_t^{\infty} e^{-rt'} p_s(\pi_s(\bar{\pi}, t), t') \left(A - \lambda_s(\pi_s(\bar{\pi}, t), t') \right) q_f^t(t') dt'. \end{aligned} \tag{3.5}$$

The first term in representation (3.1) is the value of immediate exit; the second term is the option value of waiting. See Section 8.1 for the proof.

Let $\tau > t$ denote the time of the first failure of the project if the quality is "bad." Since experimentation is not profitable after the first failure, we have $\hat{q}_f^t(t') = 0$ for all $t' > \tau$. In all the theorems below, the optimal strategies $\hat{q}_f^t(t')$ are conditioned on $t' \leq \tau$. For the brevity of exposition, we omit multiplication of the strategies by the indicator function $\mathbb{1}_{t' \leq \tau}$.

Theorem 3.2. *If $A \geq \hat{A}_s(\pi_s(\bar{\pi}, t))$, the only optimal strategy of the follower is*

$$\hat{q}_f^t(t') = 1, \ \forall \, t' > t, \tag{3.6}$$

and

$$F(\bar{\pi}, t) = S + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Phi(\bar{\pi}, t; +\infty). \tag{3.7}$$

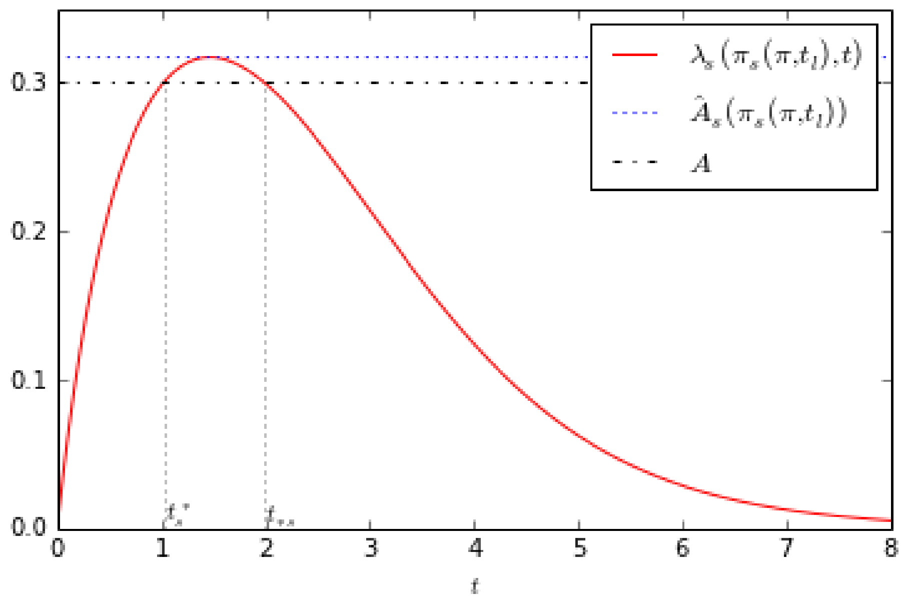

FIGURE 1. Illustration for Lemma 3.3. Parameters: $\bar{\pi} = 0.3$, $\lambda = 1$, $A = 0.3$, $t_l = 0.651$ - exit time of the leader.

Proof. Under condition $A \geq \hat{A}_s(\pi_s(\bar{\pi}, t))$, the integrand on the RHS of (3.5) is non-negative, and positive in a neighborhood of $+\infty$. Hence, the integral is maximized with the choice (3.6). □

Lemma 3.3. *Let $A < \hat{A}_s(\pi_s(\bar{\pi}, t))$. Then (a) the equation*

$$A - \lambda_s(\pi_s(\bar{\pi}, t), t') = 0 \tag{3.8}$$

has exactly two solutions $0 < t^*_s(A, \pi_s(\bar{\pi}, t)) < t_{*s}(A, \pi_s(\bar{\pi}, t))$.

(b) $t^*_s(A, \pi_s(\bar{\pi}, t))$ is the local maximum of $\Phi(A, \bar{\pi}, t; \cdot)$; $t_{*s}(A, \pi_s(\bar{\pi}, t))$ is the local minimum of $\Phi(A, \bar{\pi}, t; \cdot)$.

See Fig. 1 for illustration.

Proof. (a) The equation (3.8) can be rewritten in the form

$$at' - b = e^{\lambda t'},$$

where $a > 0$, $b < 1$. Such an equation has either two solutions, or one, or none, and any solution is positive. Under condition $A > \hat{A}_s(\pi_s(\bar{\pi}, t))$, the graphs of functions $t' \mapsto at' - b$ and $t' \mapsto e^{\lambda t'}$ do not intersect, under condition $A = \hat{A}_s(\pi_s(\bar{\pi}, t))$, the graphs are tangent, and under condition $A < \hat{A}_s(\pi_s(\bar{\pi}, t))$, the graphs of functions $t' \mapsto at' - b$ and $t \mapsto e^{\lambda t'}$ intersect at two points.
(b) If $0 < t^*_s(A, \pi_s(\bar{\pi}, t)) < t_{*s}(A, \pi_s(\bar{\pi}, t))$ are solutions to (3.8), then it is easy to see that the LHS in (3.8) is positive if $t' < t^*_s$ or $t' > t_{*s}$; and it is negative if $t^*_s < t' < t_{*s}$.

Since $\Phi(A, \bar{\pi}, t, \cdot)$ is increasing (respectively, decreasing) iff the LHS in (3.8) is positive (respectively, negative), (b) follows. □

Lemma 3.4. *For any $t \geq 0$, there exists a unique $A^*_s(\pi_s(\bar{\pi}, t)) \in \Big(0, \hat{A}_s(\pi_s(\bar{\pi}, t))\Big)$ s.t.*

$$\Phi(A, \bar{\pi}, t; +\infty) \leq 0 \ \Leftrightarrow A \leq A^*_s(\pi_s(\bar{\pi}, t)).$$

Proof. Fix $(\bar{\pi}, t)$, and suppress the dependence on $(\bar{\pi}, t)$ in the notation $\hat{A}_s$, $t^*_s(A)$, A^*_s. For $A < \hat{A}_s$, consider

$$\Phi(A, \bar{\pi}, t^*_s(A); +\infty) = C \int_{t^*_s(A)}^{\infty} e^{-rt'} p_s(\pi_s(\bar{\pi}, t), t') \left(A - \lambda_s(\pi_s(\bar{\pi}, t), t')\right) dt'. \tag{3.9}$$

The integrand in (3.9) increases in A. Furthermore, while A remains below $\hat{A}_s$, the integrand is negative in a right neighborhood of $t^*_s(A)$, and $t^*_s(A)$ moves to the right as A increases. Hence, the integral in (3.9) is increasing in A. As a function of A, the integral is positive at $\hat{A}_s$; by continuity, it is also positive in a left neighborhood of $\hat{A}_s$. In the limit $A \to +0$, the integral becomes negative. Hence, there exists $A^*_s \in (0, \hat{A}_s)$ s.t. the integral in (3.9) is negative for any $A < A^*_s$, and positive for any $A \in (A^*_s, \hat{A}_s)$. By monotonicity of $\Phi(A, \bar{\pi}, t^*_s(A); +\infty)$, this A^*_s is unique. □

Theorem 3.5. *Let $A < \hat{A}_s(\pi_s(\bar{\pi}, t))$. Then*

(1) if $A > A^*_s(\pi_s(\bar{\pi}, t))$, the follower never exits before the first failure, and the only optimal strategy is (3.6)*;*

(2) if $A < A^*_s(\pi_s(\bar{\pi}, t))$ and $t \leq t^*_s(A, \pi_s(\bar{\pi}, t))$, the follower exits at $t^*_s(A, \pi_s(\bar{\pi}, t))$. The only optimal strategy is

$$\hat{q}^t_f(t') = \begin{cases} 1, & t' \in (t, t^*_s(A, \pi_s(\bar{\pi}, t))], \\ 0, & t' > t^*_s(A, \pi_s(\bar{\pi}, t)); \end{cases} \tag{3.10}$$

(3) if $A = A^*_s(\pi_s(\bar{\pi}, t))$ and $t \leq t^*_s(A, \pi_s(\bar{\pi}, t))$, then, for any $\bar{q} \in [0, 1]$, the strategy

$$\hat{q}^t_f(t') = \begin{cases} 1, & t' \in (t, t^*_s(A, \pi_s(\bar{\pi}, t))], \\ \bar{q}, & t' > t^*_s(A, \pi_s(\bar{\pi}, t)), \end{cases} \tag{3.11}$$

is optimal, and any optimal strategy is of the form (3.11)*;*

(4) if $A < A^*_s(\pi_s(\bar{\pi}, t))$ and $t > t^*_s(A, \pi_s(\bar{\pi}, t))$, then the only optimal strategy is

$$\hat{q}^t_f(t') = 0, \ t' > t; \tag{3.12}$$

(5) if $A = A^*_s(\pi_s(\bar{\pi}, t))$ and $t > t^*_s(A, \pi_s(\bar{\pi}, t))$, then, for any $\bar{q} \in [0, 1]$, the strategy

$$\hat{q}^t_f(t') = \bar{q}, \ \forall t' > t, \tag{3.13}$$

is optimal, and any optimal strategy is of the form (3.13).

Obviously, cases (3) and (5) are non-generic, while the rest of the cases in Theorem 3.5 are generic ones. All statements of the theorem are immediate from the following Lemma.

Lemma 3.6. *Let the following conditions hold*

(i) F_1 *is a piece-wise continuous function and* $\exists B > 0$ *and* $r > 0$ *s.t.* $|F_1(t)| \le Be^{-rt}\ \forall\, t \ge 0$.

(ii) Function F_2 *defined by*

$$F_2(T) = \int_t^T F_1(t')dt'$$

has a finite number $t \le t_1 < t_2 < \cdots t_n \le +\infty$ *of points of the global maximum.*

(iii) $q : [t, +\infty) \to [0, 1]$ *is a non-decreasing LCRL function with the trivial singular continuous component, s.t.* $q(t) = 1$;

Then the problem

$$V = \sup_q \int_t^{+\infty} F_1(t')q(t')dt'$$

where the supremum is taken over the class of function satisfying (iii), has solutions of the form

$$\hat{q}_f^t(t') = \begin{cases} 1, & t' \in [t, t_1], \\ \bar{q}_j, & t' \in (t_j, t_{j+1}], j = 1, 2, \dots, n-1, \\ 0, & t' > t_n, \end{cases} \tag{3.14}$$

where $1 \ge \bar{q}_1 \ge \bar{q}_2 \ge \cdots \ge \bar{q}_n \ge 0$, *and any optimal solution is of the form* (3.14).

Proof. Integrating by parts, we obtain

$$\begin{aligned} V &= -F_2(t) + F_2(+\infty)q(+\infty) + \int_t^{+\infty} F_2(t')(-dq(t')) \\ &= F_2(+\infty)q(+\infty) + \int_t^{+\infty} F_2(t')(-dq(t')). \end{aligned}$$

and the statement follows. □

The critical value A_s^* is s.t. for all $A < A_s^*$, $t_s^*(A)$ is the global maximum of $\Phi(A, \bar{\pi}, t; \cdot)$. If $A > A_s^*$, the global maximum of $\Phi(A, \bar{\pi}, t; \cdot)$ is at $T = +\infty$. If $A = A_s^*$, $\Phi(A, \bar{\pi}, t; \cdot)$ (which is a non-generic case) has two maxima - $T = t_s^*(A)$ and $T = +\infty$. The statements of Theorems 3.2 and 3.5 imply that, in a generic case, the follower may find it optimal to exit at the same time as the leader, never to exit before the first failure, or exit some time after the leader's exit unless the first failure happens earlier. Let $t_f = t_f(A, \bar{\pi}; t)$ denote the optimal stopping time of the follower, which may be t, $t_s^*(A, \pi_s(\bar\pi, t))$, or $+\infty$. Then, in a generic case, we can write the follower's optimal strategy as

$$\hat q_f^t(t') = \begin{cases} 1, & t' \in (t, t_f], \\ 0, & t' > t_f; \end{cases}$$

and the follower's value as

$$F(\bar\pi, t) = S + \frac{e^{rt}}{p_d(\bar\pi, t)} \Phi(A, \bar\pi, t; t_f). \tag{3.15}$$

Corollary 3.7. *If $t < t_s^*(A, \pi_s(\bar\pi, t))$ and $A < A_s^*(\pi_s(\bar\pi, t))$, the follower's optimal strategy is*

$$\hat q_f^t(t') = \begin{cases} 1, & t' \in (t, t_s^*(A, \pi_s(\bar\pi, t))], \\ 0, & t' > t_s^*(A, \pi_s(\bar\pi, t)); \end{cases}$$

and

$$F(\bar\pi, t) = S + \frac{e^{rt}}{p_d(\bar\pi, t)} \Phi(A, \bar\pi, t; t_s^*(A, \pi_s(\bar\pi, t))). \tag{3.16}$$

Hence, the follower exits at time $t_s^*(A, \pi_s(\bar\pi, t))$ unless the first failure happens earlier.

Recall that $t_s^*(A, \pi_s(\bar\pi, t))$ is the first time the rate of arrival $\lambda_s(\bar\pi, t; t')$ of the time-inhomogeneous Poisson process defined by the "survival" probability $p_s(\bar\pi, t; t')$ reaches its critical value A. In the literature dealing with two-armed Poisson bandits, it is common to formulate stoping strategies in terms of critical beliefs about the quality of the "risky" arm. In particular, in the case of failures, the beliefs about the project being "bad" and the rate of arrival of the corresponding time-inhomogeneous Poisson process are decreasing functions of time. To be more specific, these two functions differ only by a positive factor. That is why, in that model, it is either optimal to stop immediately, or never to stop before the first failure, and formulation of the optimal stopping strategy in terms of the critical beliefs or the critical rate of arrival is equivalent.

Comparing (2.5) and (2.7), we see that

$$\lambda_s(\bar\pi, t; t') = \frac{\lambda t'}{1 + \lambda t'} (1 - \pi_s(\bar\pi, t; t')),$$

i.e., the rate of arrival is the product of $1 - \pi_s(\bar\pi, t; t')$, the beliefs about the project being "bad", which decreases in time, and the rate of arrival $\lambda^2 t'/(1+\lambda t')$ of the time-inhomogeneous Poisson process defined by the "survival" probability $p_\lambda(s')$, which increases in time. Hence, $\lambda_s(\bar\pi, t; t')$ and $1 - \pi_s(\bar\pi, t; t')$ may behave differently as functions of time, and formulation of the optimal stopping strategy in terms of the critical rate of arrival is more adequate than in terms of critical beliefs.

Also, notice that the critical level $\lambda_s(\bar\pi, t; t')$ is achieved, when the marginal benefit of staying active $r(R - S)$ reaches the marginal expected cost $\lambda_s(\bar\pi, t; t')C$ for the first time. If the net expected marginal benefit $r(R - S) - \lambda_s(\bar\pi, t; t')C$ is positive for all $t' \geq t$, it is never optimal to exit before the first failure ($t_f(A, \pi_s(\bar{\pi}, t) = \infty)$. If the net expected marginal benefit is negative for t' is a right neighborhood of t, then it is necessary to calculate the net expected life-time benefit and exit either immediately ($t_f(A, \pi_s(\bar{\pi}, t) = t)$ if the net expected life-time benefit is negative) or never ($t_f(A, \pi_s(\bar{\pi}, t) = +\infty)$ if the net expected life-time benefit is positive).

3.2. **Value of player i.** Consider a subgame starting at $t \geq 0$ after a history such that none of the players has yet acted. Consider the value function of player i in such a subgame. We have

$$\begin{aligned} V_i(\bar{\pi}, t; q_i^t, q_j^t) \;\; = \;\; & \int_t^{\infty} e^{-r(t'-t)} \frac{p_d(\bar{\pi}, t')}{p_d(\bar{\pi}, t)} G_i(\bar{\pi}, t') q_i^t(t') q_j^t(t') dt' \qquad (3.17) \\ & + \int_{\{t' \geq t \;|\; \Delta q_i^t(t')=0\}}^{\infty} e^{-r(t'-t)} \frac{p_d(\bar{\pi}, t')}{p_d(\bar{\pi}, t)} F_i(\bar{\pi}, t') q_i^t(t')(-dq_j^t(t')) \\ & + \int_t^{\infty} e^{-r(t'-t)} \frac{p_d(\bar{\pi}, t')}{p_d(\bar{\pi}, t)} S q_j^t(t')(-dq_i^t(t')), \end{aligned}$$

where, as before, $G_i(\bar{\pi}, t')$ is the expected payoff flow that player i gets when both players are experimenting. The flow $G_i(\bar{\pi}, t')$ has several components. The revenue flow is rR; if a failure happens to one of the players, then this player pays the cost of failure $\hat{C}$, but both players suffer the loss L in the recovery value. Assuming that, if the project is bad, the players are equally likely to incur a costly failure, we can write

$$G_i(\bar{\pi}, t') = rR + \lambda_d(\bar{\pi}, t')(S - L - 0.5\hat{C}).$$

Recall that we set $C = L + \hat{C}$, which is the total loss in case of a failure. Therefore, the expected payoff in case of a failure can be written as $S - 0.5(L + C)$. Introduce $\kappa = (L + C)/C$, then $\kappa \in [1, 2)$, where $\kappa = 1$ corresponds to the case $L = 0$ - no loss in the recovery value; $\kappa \to 2$ as $\hat{C} \to 0$. Given this notation, we rewrite

$$G_i(\bar{\pi}, t') = rR + \lambda_d(\bar{\pi}, t')(S - 0.5\kappa C).$$

Lemma 3.8. *We have*

$$\begin{aligned} V_i(\bar\pi, t; q_i^t, q_j^t) &= S \\ &+ \frac{e^{rt}}{p_d(\bar\pi,t)} \int_t^\infty \Big[C e^{-rt'} p_d(\bar\pi,t') \left(A - 0.5\kappa\lambda_d(\bar\pi,t')\right) q_j^t(t') \\ &\qquad - \Phi(A,\bar\pi,t'; t_f(A,\bar\pi,t'))(q_j^t(t'))' \Big] q_i^t(t') dt \\ &+ \frac{e^{rt}}{p_d(\bar\pi,t)} \sum_{\substack{t' \ge t: \\ \Delta q_i(t') = 0 \\ \Delta q_j(t') \neq 0}} \Phi(A,\bar\pi,t'; t_f(A,\bar\pi,t')) q_i^t(t') (-\Delta q_j(t')) \\ &- \frac{e^{rt}}{p_d(\bar\pi,t)} \sum_{\substack{t' \ge t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} S e^{-rt'} \left(q_i^t(t') + \Delta q_i^t(t') \right) (-\Delta q_j(t')). \end{aligned} \tag{3.18}$$

See Section 8.2 for the proof.

4. Equilibria

4.1. SPE, where both players stay until the first failure.

Let

$$\hat{A}_d = \hat{A}_d(\kappa, \bar\pi) = 0.5\kappa \max_{t \ge 0} \lambda_d(\bar\pi, t), \tag{4.1}$$

and

$$\Psi(\kappa, A, \bar\pi, t; T) = C \int_t^T e^{-rt'} p_d(\bar\pi, t') \left(A - 0.5\kappa\lambda_d(\bar\pi,t')\right) dt'. \tag{4.2}$$

Let $\tau > t$ denote the time of the first failure of the project if the quality is "bad." Since experimentation is not profitable after the first failure, we have $\hat{q}_i^t(t') = \hat{q}_j^t(t') = 0$ for all $t' > \tau$. In all the theorems below, the optimal strategies $\hat{q}_i^t(t')$ ($i \in \{1,2\}$) are conditioned on $t' \le \tau$. For the brevity of exposition we omit multiplication of strategies by the indicator function $\mathbb{1}_{t' \le \tau}$.

Theorem 4.1. *If $A \ge \hat{A}_d(\kappa, \bar\pi)$, then for any $t \ge 0$, the following profile is a SPE in the subgame starting at t: for $i, j \in \{1,2\}$, $i \neq j$:*

$$\hat{q}_i^t(t') = \hat{q}_1^t(t') = 1, \quad t' > t, \tag{4.3}$$

and

$$V_i(\bar\pi, t; \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar\pi, t; \hat{q}_i^t, \hat{q}_j^t) = S + \frac{e^{rt}}{p_d(\bar\pi,t)} \Psi(\kappa, A, \bar\pi, t; +\infty). \tag{4.4}$$

See Section 8.3 for the proof. Theorem 4.1 states that if $A = r(R-S)/C$ is sufficiently large, then in a SPE, the players stop simultaneously at the moment of the first failure if the project is "bad," or never if the project is "good."

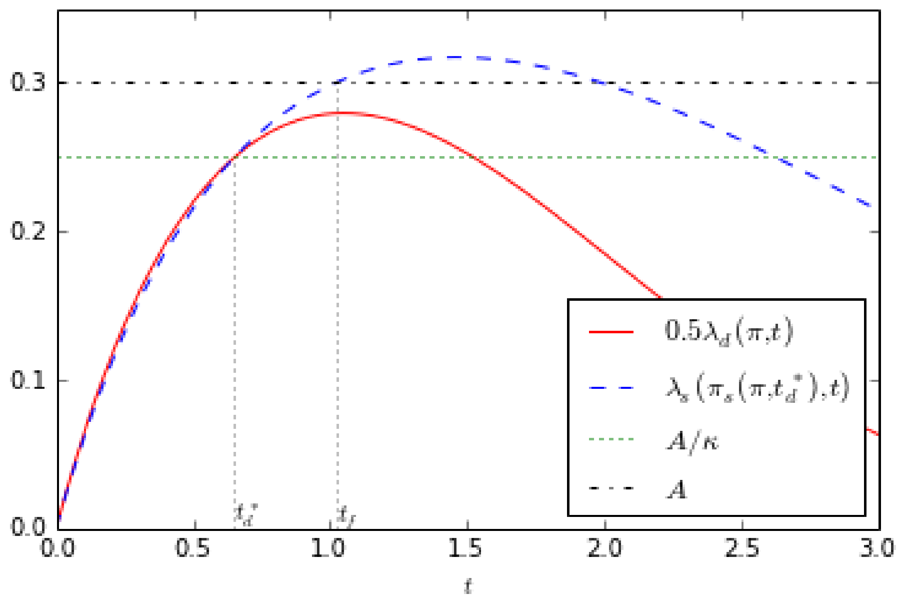

FIGURE 2. Illustration for Lemma 4.2. Parameters: $\bar{\pi} = 0.3$, $\lambda = 1$, $A = 0.3$ $\kappa = 1.2$.

Lemma 4.2. *Let* $A < \hat{A}_d(\kappa, \bar{\pi})$, *then*

a) the equation

$$A - 0.5\kappa\lambda_d(\bar{\pi}, t) = 0 \tag{4.5}$$

has two solutions $t_d^*(\kappa, A, \bar{\pi}) < t_{*d}(\kappa, A, \bar{\pi})$ *s.t.* $t_d^*(\kappa, A, \bar{\pi})$ *is the local maximum, and* $t_{*d}(\kappa, A, \bar{\pi})$ *is the local minimum of* $\Psi(A, \bar{\pi}, t; \cdot)$.

b) $t_d^*(\kappa, A, \bar{\pi}) \le t_f(\kappa, A, \pi_s(\bar{\pi}, t_d^*))$ *for all* $\kappa \in [1, 2)$ *and* $A < \hat{A}_d(\kappa, \bar{\pi})$, *and the equality holds only if* $\kappa = 1$.

See Fig. 2 for illustration. See Section 8.4 for the proof.

Remark 4.3. Lemma 4.2 implies that

(i) In any subgame that starts at time $t < t_d^* = t_d^*(\kappa, A, \bar{\pi})$ (we suppress the dependence of t_d^* on $(\kappa, A, \bar{\pi})$ in order to simplify the notation) after a history such that none of the players has yet acted, none of the players will find it optimal to exit before t_d^*, because

$$\psi(A, \bar{\pi}, t') := A - 0.5\kappa\lambda_d(\bar{\pi}, t') \tag{4.6}$$

is positive for all $t \le t' < t_d^*$. To understand why, recall that

$$C\psi(A, \bar{\pi}, t') = r(R - S) - 0.5\kappa\lambda_d(\bar{\pi}, t')C$$

is the net expected marginal benefit of each player, when two players keep experimenting. It is not optimal to stop while the net marginal benefit is positive.

(ii) If one of the players (the leader) decides to exit at time $t_d^* = t_d^*(\kappa, A, \bar{\pi})$, the other player (the follower) will find it optimal to stay longer (may be until the first failure if $t_f(A, \pi_s(\bar{\pi}, t_d^*)) = \infty$). The only case when the follower finds it optimal to exit together with the leader at $t_d^*(\kappa, A, \bar{\pi})$ is when $\kappa = 1$, i.e., when there is no loss in the recovery value due to a failure. To see why, recall that if a player exits the game in finite time, the optimal stopping rule can be formulated in terms of the critical value of the corresponding rate of arrival (i.e. the point where the expected net marginal benefit is zero). If there is only one player, this critical level is A, if there are two players, this critical value is $2A/\kappa$. If $\kappa = 1$, the corresponding critical levels of rates of arrival are achieved simultaneously. If $\kappa > 1$, the critical rate of arrival for two players is achieved earlier than the critical level for a single agent (the follower). That is why the follower delays his/her exit if the leader exits at $t_d^*(\kappa, A, \bar{\pi})$.

We have the following analog of Lemma 3.4.

Lemma 4.4. *There exists a unique* $A_d^* = A_d^*(\kappa, \bar{\pi}) \in (0, \hat{A}_d(\kappa, \bar{\pi}))$ *s.t.*

$$\Psi(\kappa, A_d^*, \bar{\pi}, t_d^*; +\infty) = 0, \tag{4.7}$$

and $\Psi(\kappa, A, \bar{\pi}, t_d^*; +\infty) > 0 (= 0, < 0)$ *if* $A > A_d^* (= A_d^*, < A_d^*)$.

Proof follows the proof of Lemma 3.4 line by line.

Theorem 4.5. *Let* $A \geq A_d^*(\kappa, \bar{\pi})$*, then for* $t \geq 0$, $(\hat{q}_i^t, \hat{q}_j^t)$ *given by* (4.3) *is a SPE, and the payoffs are given by* (4.4).

Proof. If $A > A_d^* = A_d^*(\kappa, \bar{\pi})$, the function $\Psi(\kappa, A, \bar{\pi}, t, T)$ is maximized at $T = +\infty$. Hence if player j plays $\hat{q}_j^t$, the best response of player i is to play $\hat{q}_i^t$ and vice versa. By definition, $\Psi(A_d^*, \bar{\pi}, t_d^*; +\infty) = 0$, hence for any $t > t_d^*$, $\Psi(A_d^*, \bar{\pi}, t; +\infty) > 0$, hence it is never optimal to exit before the first failure happens, hence $(\hat{q}_i^t, \hat{q}_j^t)$ given by (4.3) is a SPE. If $t \leq t_d^*$, it is never optimal to exit earlier than at t_d^*, because $\Psi(\kappa, A, \bar{\pi}, t, T)$ is increasing in T if $T \in (t, t_d^*]$. Since $\Psi(\kappa, A_d^*, \bar{\pi}, t_d^*; +\infty) = 0$, the players are indifferent between exiting at t_d^* and staying until the first failure happens as long as they exit or stay together, but exiting at t_d^* simultaneously is not an equilibrium, because if one of the players exits at that time, the other player has an incentive to wait until $t_f(A, \pi_s(\bar{\pi}, t_d^*))$ is reached. □

4.2. **Symmetric SPE, where players stop before the first failure.** Let $A < A_d^*(\kappa, \bar{\pi})$, and let $\hat{T}_d = \hat{T}_d(\kappa, A, \bar{\pi}) > t_d^*(\kappa, A, \bar{\pi})$ be the solution to

$$\Psi(\kappa, A, \bar{\pi}, \hat{T}_d, +\infty) = 0.$$

Then, for $t < \hat{T}_d$, $\Psi(\kappa, A, \bar{\pi}, t, +\infty) < 0$, and for $t > \hat{T}_d$, $\Psi(\kappa, A, \bar{\pi}, t, +\infty) > 0$.

Theorem 4.6. *Consider a subgame that starts at* $t \geq \hat{T}_d(\kappa, A, \bar{\pi}))$ *after a history such that none of the players has yet acted, then for all* $t \geq \hat{T}_d(\kappa, A, \bar{\pi}))$, $(\hat{q}_i^t, \hat{q}_j^t)$ *given by* (4.3) *is a SPE, and the payoffs are given by* (4.4).

Proof is the same as the proof of Theorem 4.5

Let $A < A^*_d(\kappa,\bar\pi)$, $\kappa > 1$, and let $\hat T_f(A,\bar\pi;t) = \inf\{t' \ge t \mid t_f(A,\pi_s(\bar\pi,t')) = t'\}$.

Lemma 4.7. *$\hat T_f(A,\bar\pi;t)$ is independent of $t > t^*_d(\kappa,A,\bar\pi)$ such that $t_f(A,\pi_s(\bar\pi,t)) > t$.*

Proof. If $t_f(A,\pi_s(\bar\pi,t')) > t'$ for all $t' > t$, then $\hat T_f(A,\bar\pi;t') = +\infty$ for all $t' \ge t$. Otherwise, $\hat T_f(A,\bar\pi;t')$ is the first zero above t of the function $t' \mapsto t_f(A,\pi_s(\bar\pi,t')) - t'$. □

Thus, we may write $\hat T_f(A,\bar\pi)$ rather than $\hat T_f(A,\bar\pi;t)$. If $\hat T_f(A,\bar\pi) < \hat T_d(\kappa,A,\bar\pi)$, then in the interval $[\hat T_f(A,\bar\pi), \hat T_d(\kappa,A,\bar\pi))$, experimentation is not optimal either for one or for two players, so for any subgame that starts at t is the latter interval, the SPE equilibrium is $\hat q^t_i(t') = \hat q^t_j(t') = 0$, for all $t' > t$.

Consider the case when $\hat T_f(A,\bar\pi) \le \hat T_d(\kappa,A,\bar\pi)$. Introduce functions

$$\begin{aligned} U_0(t') &:= U_0(A,\bar\pi,t') = e^{-rt'}p_d(\bar\pi,t')C\psi(A,\bar\pi,t') \\ U_1(t') &:= \Phi(A,\bar\pi,t',t_f(A,\pi_s(\bar\pi,t'))), \end{aligned}$$

and consider the Cauchy problem

$$U_0(t')\alpha(t,t') - U_1(t')\frac{\partial\alpha}{\partial t'}(t,t') = 0, \tag{4.8}$$

subject to $\alpha(t,t) = 1$, with t as a parameter.

Lemma 4.8. *a) Function $\alpha(t,\cdot)$ is well-defined on $[t,\hat T_f)$.*

b) As $t' \to \hat T_f$, $\alpha(t,t') \downarrow 0$.

c) For $t < t_1 < t' < T_f$, $\alpha(t,t_1)\alpha(t_1,t') = \alpha(t,t')$

Proof. a) We have $\alpha(t,t') = \exp(I(t,t'))$, where

$$I(t,t') = \int_t^{t'} \frac{U_0(y)}{U_1(y)}dy.$$

b) Function U_0 is continuous and negative on $(t,\hat T_f]$; U_1 is positive on $(t,\hat T_f)$ and $U_1(T_f) = 0$. Furthermore, using the implicit function theorem, we obtain that $t' \mapsto t_f(A,\pi_s(\bar\pi,t'))$ is differentiable, hence, U_1 is differentiable as well. We conclude that $0 < U_1(t') < C(T_f - t')$, where $C > 0$ is independent of $t' \in [t,\hat T_f)$. Hence, as $t' \to \hat T_f$, $I(t,t') \to -\infty$, and $\alpha(t,t') \downarrow 0$.

c) It suffices to note that

$$I(t,t') = \int_t^{t'} \frac{U_0(y)}{U_1(y)}dy = \left(\int_t^{t_1} + \int_{t_1}^{t'}\right)\frac{U_0(y)}{U_1(y)}dy = I(t,t_1) + I(t_1,t').$$

□

Theorem 4.9. *Let $A < A_d^*(\kappa, \bar{\pi})$, $\kappa > 1$, and $\hat{T}_f(A, \bar{\pi}) \le \hat{T}_d(\kappa, A, \bar{\pi})$. Consider a subgame that starts at $t \in [t_d^*(\kappa, A, \bar{\pi}), \hat{T}_f(A, \bar{\pi}))$ after a history such that none of the players has yet acted. Then, there exists a symmetric SPE given by the following pair of simple consistent strategies:*

$$\hat{q}_j^t(t') = \hat{q}_j^t(t') = \begin{cases} \alpha(t, t'), & t' \in (t, \hat{T}_f(A, \bar{\pi})] \\ 0, & t' \ge \hat{T}_f(A, \bar{\pi}). \end{cases} \tag{4.9}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = S. \tag{4.10}$$

Proof. We apply Lemma 3.8. If agent j follows the strategy (4.9), then 1) there are no jumps Δq_j^t, and all the jump terms on the RHS of (3.18) are 0; 2) the integral on the RHS of (3.18) is 0 for any choice of q_i^t. Hence, any deviation gives the same value S. □

Theorem 4.10. *Let $A < A_d^*(\kappa, \bar{\pi})$, $\kappa > 1$, and $\hat{T}_f(A, \bar{\pi}) \le \hat{T}_d(\kappa, A, \bar{\pi})$. Consider a subgame that starts at $t \le t_d^*(\kappa, A, \bar{\pi})$ after a history such that none of the players has yet acted Then, there exists a symmetric SPE given by the following pair of simple consistent strategies:*

$$\hat{q}_j^t(t') = \hat{q}_j^t(t') = \begin{cases} 1, & t' \le t_d^*(\kappa, A, \bar{\pi}) \\ \alpha(t_d^*(\kappa, A, \bar{\pi}), t'), & t' \in (t_d^*(\kappa, A, \bar{\pi}), \hat{T}_f(A, \bar{\pi})) \\ 0, & t' \ge \hat{T}_f(A, \bar{\pi}). \end{cases} \tag{4.11}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = S + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d^*). \tag{4.12}$$

Proof. We apply Lemma 3.8. If agent j follows the strategy (4.11), then 1) there are no jumps Δq_j^t, and all the jump terms on the RHS of (3.18) are 0; 2) the integral in the RHS of (3.18) can be written as

$$\Psi(\kappa, A, \bar{\pi}, t; t_d^*) + \Psi(\kappa, A, \bar{\pi}, t_d^*, \hat{T}_f),$$

where the second term is 0 for any choice of q_i^t. Hence, any deviation gives the same value as in (4.12). □

Remark 4.11. a) Since $\alpha(t, \hat{T}_f(A, \bar{\pi}) - 0) = 0$, functions $q_i^t = q_j^t$ are continuous.

b) The values at t_d^* are the same as if both players exit at $t = t_d^*$ but the simultaneous exit at t_d^* is not an equilibrium.

c) The symmetric equilibrium is inefficient: it is dominated by each of the asymmetric equilibria that we characterize in the next Section.

It remains to mention a special case, when the symmetric SPE is efficient.

Theorem 4.12. *Let $A < A_d^*(\kappa,\bar\pi)$, $\kappa = 1$. Consider a subgame that starts at $t \le t_d^*(\kappa, A,\bar\pi)$ after a history such that none of the players has yet acted Then, there exists a symmetric SPE given by the following pair of simple consistent strategies:*

$$\hat q_j^t(t') = \hat q_j^t(t') = \begin{cases} 1, & t' \le t_d^*(\kappa, A, \bar\pi) \\ 0, & t' \ge t_d^*(\kappa, A, \bar\pi). \end{cases} \tag{4.13}$$

The players payoffs are

$$V_i(\bar\pi, t; \hat q_i^t, \hat q_j^t) = V_j(\bar\pi, t; \hat q_i^t, \hat q_j^t) = S + \frac{e^{rt}}{p_d(\bar\pi, t)}\Psi(\kappa, A, \bar\pi, t; t_d^*). \tag{4.14}$$

If $\kappa = 1$, then $t_d^*(1, A, \bar\pi) = t_f(A, \pi_s(\bar\pi, t_d^*))$, hence simultaneous stopping is a SPE.

4.3. **Asymmetric equilibria with precommitment.**

Theorem 4.13. *Let $A < A_d^*(\kappa,\bar\pi)$, $\kappa > 1$, and $\hat T_f(A,\bar\pi) \le \hat T_d(\kappa, A, \bar\pi)$. Consider a subgame that starts at $t \in [t_d^*(\kappa, A, \bar\pi), \hat T_f(A, \bar\pi))$ after a history such that none of the players has yet acted. Then there are two asymmetric equilibria given by the following pairs of simple consistent strategies: for $(i, j) \in \{(1, 2), (2, 1)\}$,*

$$\hat q_i^t(t') = 0,\ \forall\, t' > t, \tag{4.15}$$

and

$$\hat q_j^t(t') = \begin{cases} 1, & t' \in (t, \hat T_f(A,\bar\pi)], \\ 0, & t' > \hat T_f(A, \bar\pi). \end{cases} \tag{4.16}$$

The players payoffs are

$$V_i(\bar\pi, t; \hat q_i^t, \hat q_j^t) = S, \quad V_j(\bar\pi, t; \hat q_i^t, \hat q_j^t) = F(\bar\pi, t).$$

Proof. Follows immediately from the fact that experimentation is not optimal for two players, but optimal for one player if $t \in [t_d^*(\kappa, A, \bar\pi), \hat T_f(A, \bar\pi))$. □

Corollary 4.14. *If the same type of equilibrium is played in any subgame that starts at $t \in [t_d^*(\kappa, A, \bar\pi), \hat T_f(A, \bar\pi))$, then the strategy profile* (4.15)-(4.16) *is a SPE.*

Indeed, if player j precommits not to stop in any subgame that starts at $t \in [t_d^*(\kappa, A, \bar\pi), \hat T_f(A, \bar\pi))$, then player i's best response is to stop first. On the other hand, if player i will be the first one to stop, then player j's best response is to be the follower.

Theorem 4.15. *Let $A < A_d^*(\kappa,\bar\pi)$, $\kappa > 1$, and $\hat T_f(A,\bar\pi) \le \hat T_d(\kappa, A, \bar\pi)$. In a subgame that starts at $0 \le t < t_d^* = t_d^*(\kappa, A, \bar\pi)$ after a history such that none of the players has yet acted, there are two asymmetric equilibria given by the following pairs of simple consistent strategies: for $(i, j) \in \{(1, 2), (2, 1)\}$,*

$$\hat q_i^t(t') = \begin{cases} 1, & t' \in (t, t_d^*], \\ 0, & t' > t_d^*; \end{cases} \tag{4.17}$$

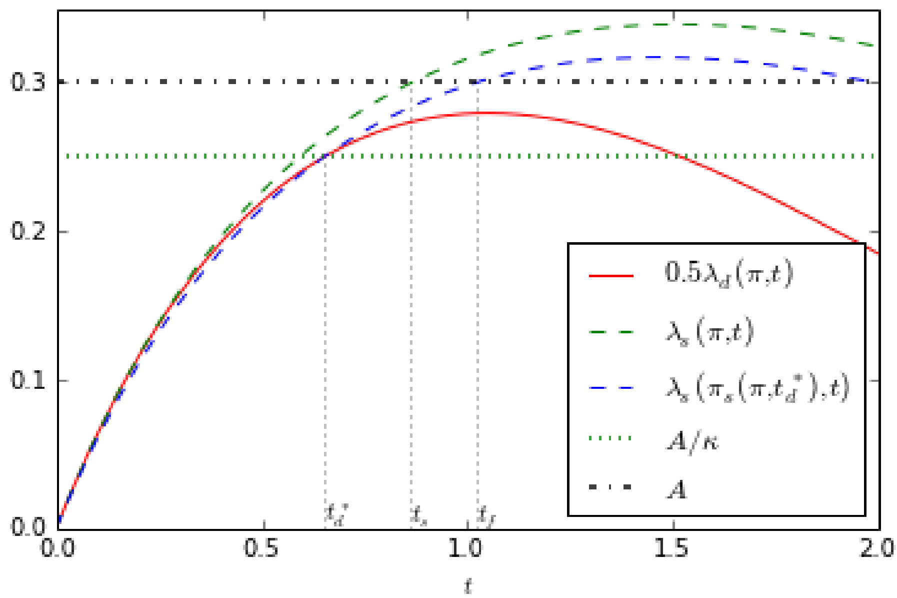

FIGURE 3. Illustration for Lemma 4.17. Parameters: $\bar{\pi} = 0.3$, $\lambda = 1$, $A = 0.3$ $\kappa = 1.2$.

and

$$\hat{q}_j^t(t') = \begin{cases} 1, & t' \in (t, t_f(t_d^*)], \\ 0, & t' > t_f(t_d^*). \end{cases} \tag{4.18}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = S + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d^*), \tag{4.19}$$

$$V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = F(\bar{\pi}, t_d^*) + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d^*). \tag{4.20}$$

Proof. Let player i play the strategy (4.17), then the best response of player j is the strategy (4.18) as the optimal follower's strategy that we derived in Section 3.1. Let player j play the strategy (4.18), then $\Psi(\kappa, A, \bar{\pi}, t, T)$ is maximized at $t = t_d^*(\kappa, A, \bar{\pi})$. By Lemma 3.6, the strategy (4.17) is the optimal strategy of player i.

□

Corollary 4.16. *If $\hat{T}_f(A, \bar{\pi}) \leq \hat{T}_d(\kappa, A, \bar{\pi})$ and the same type of equilibrium is played in any subgame that starts at $t \in [t_d^*(\kappa, A, \bar{\pi}), \hat{T}_f(A, \bar{\pi}))$, then the strategy profile (4.17)-(4.18) is a SPE.*

An advantage of being an ostrich. We finish the section with the study of the role of information externality in the stopping game with conclusive failures. To this end,

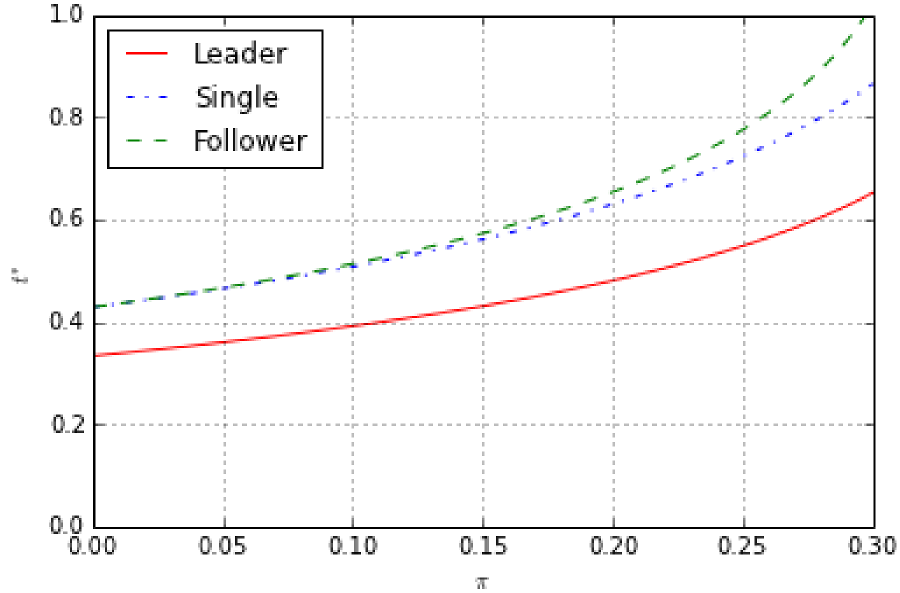

FIGURE 4. Dependence of optimal stopping times on $\bar{\pi}$. Parameters: $\lambda = 1$, $A = 0.3$ $\kappa = 1.2$.

consider a single experimenter with the value function

$$V_s(\bar{\pi}) = S + C \sup_{q_s} \int_0^\infty e^{-rt} p_s(\bar{\pi}, t) \left(A - \lambda_s(\bar{\pi}, t)\right) q_s(t) dt.$$

Let $t_s = t_s(A, \bar{\pi})$ be the single player's optimal stopping time, which may be finite or infinite. If $t_s(A, \bar{\pi}) < +\infty$, then it is the smallest solution to

$$A - \lambda_s(\bar{\pi}, t) = 0. \tag{4.21}$$

Lemma 4.17. *For any $\bar{\pi} \in (0, 1)$, there exists $\kappa^*(\bar{\pi}) \in (1, 2)$ such that for any $\kappa \in (\kappa^*(\bar{\pi}), 2)$ and any $A < A^*_d(\bar{\pi}, \kappa)$, $t^*_d(\kappa, A, \bar{\pi}) < t_s(A, \bar{\pi}) < t_f(A, \pi_s(\bar{\pi}, t^*_d))$, and $V_s(\bar{\pi})$ is higher than the value of the leader in Theorem 4.15.*

See Section 8.5 for the proof. See Fig. 3 for illustration. In Fig. 4, we show dependence of the optimal stopping times on $\bar{\pi}$, and in Fig. 4 - dependence of value functions on $\bar{\pi}$.

5. Conclusive breakthroughs

5.1. **The setup.** In this Section, we consider the game of timing, characterized by the following structure. Time $t \in \mathbb{R}_+$ is continuous, and the discount rate is $r > 0$. Two symmetric players experiment with technologies of unknown quality. The quality of a project depends on the state of nature $\theta \in \{0, 1\}$. If $\theta = 1$, the project is "good," which means that it generates positive revenues (breakthroughs). If $\theta = 0$,

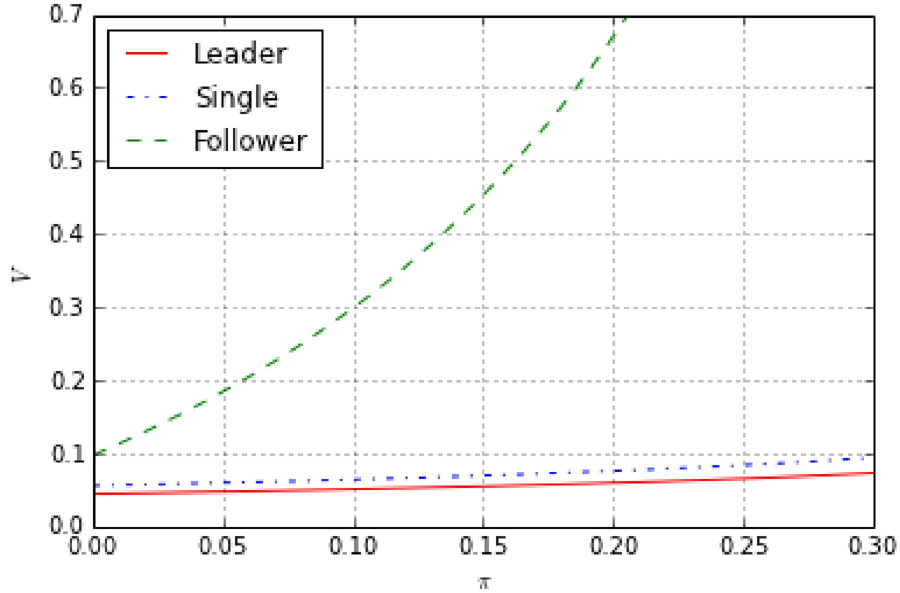

FIGURE 5. Dependence of value functions on $\bar{\pi}$. Parameters: $\lambda = 1$, $R = 30$, $S = 0$, $C = 1$ $\kappa = 1.2$, $r = 0.01$.

the project is "bad," which means that the player experimenting with this project will never be able to generate positive revenues. Experimentation is costly, and the stream of experimentation costs is rC for each player, independent of the quality of the project. The initial common prior assigns probability $\bar{\pi} \in (0, 1)$ to $\theta = 0$.

Let the quality of the project be the same for both players, so that if one of the players observes a breakthrough, both players know that their technologies are "good." For simplicity, assume that if the project if "good", both players stop experimentation after the first success had been observed (provided that both are active at that moment). Further assume that, if the project is "good", the player who is first to succeed gets $R_1 > 0$, and the other player gets $R_2 \in (0, R_1)$. Thus, there is an advantage to generating the success first (though, in this stylized model, this advantage is independent of the players' actions), which can be interpreted as an opportunity to file a patent. The other player can sell the research facility for R_2, just knowing that the project is good, but it would be necessary to pay for the right to use it. If a player exits before the first observation of a success, then the recovery value is $R_3 \in (0, (R_1 + R_2)/2)$ - i.e., without any "good" news, the player can still sell the research facility, but since the quality of the project remains unknown the value of the outside option is less than the expected value at the time of the observation of the first success.

Let τ_i denote the random time of the first breakthrough of player i if $\theta = 1$. We assume that τ_i and τ_j are i.i.d. $Erlang(2, \lambda)$ random variables. Note that the expected time until the first breakthrough is $2/\lambda$.

In the current setting, we consider the case when all payoffs, parameters of the Erlang distribution, and the players' actions are public information. W.l.o.g. assume that the game starts at $t = 0$. At each point $t \geq 0$, player $i \in \{1, 2\}$ may make an irreversible stopping decision conditioned on the history of the game. At any $t \geq 0$, the history of the game includes observations of all failures (including the empty set if no failures were observed by the players up to time t) and the actions of the players. As far as the actions are concerned, only two sorts of histories matter in the stopping game: (i) both players are still in the game; (ii) at least one player exited the game.

Let $T_i \in \mathbb{R}_+$ denote the exit time of player i. Define the function

$$\tilde{t}_i(t) = \begin{cases} T_i, & \text{if } T_i \leq t, \\ \infty, & \text{otherwise.} \end{cases}$$

Let τ_i^s denote a random time, when a breakthrough in player i's project occurred for the s^{th} time. The history of observations at any $t \geq 0$ is

$$O_t = \{(\tau_1^{s'})\}_{s' \leq t \wedge T_1} \cup \{(\tau_2^{s''})\}_{s'' \leq t \wedge T_2}.$$

If $O_t = \{\emptyset\}$, then a typical history at time t is $h_t(O_t, \tilde{t}_1(t), \tilde{t}_2(t))$. If $T_i < T_j$, we call player i the leader, and player j the follower. If $O_t \neq \{\emptyset\}$, the game is over.

From Section 2.1, we borrow Definitions 2.1, 2.2 of simple strategies and Definition 2.3 of consistent strategies. The evolution of beliefs, "survival" probabilities and rates of arrival are given by the corresponding equations in Section 2.2.

5.2. **Value functions and equilibrium.** Consider the game that starts at $t = 0$. Assume that

$$\sup_{T>0} \left(\int_0^T e^{-rt} p_d(\bar{\pi}, t') \left(\lambda_d(\bar{\pi}, t') \frac{R_1 + R_2}{2} - rC \right) dt' + e^{-rT} p_d(\bar{\pi}, T) R_3 \right) > R_3. \quad (5.1)$$

Then it is optimal for both players to experiment until the expected rate of arrival reaches its maximal value

$$\hat{t}_d(\bar{\pi}) = \arg\max_{t \geq 0} \lambda_d(\bar{\pi}, t),$$

and for some time after that. We leave for the future work the analysis of different scenarios, when given parameters of the model, it may be optimal to start experimentation for two players, but not for one, and vice versa. As the rate of arrival $\lambda_d(\bar{\pi}, t)$ starts decreasing, it may become optimal for one or both players to quit. We will show that the stopping rules in pure strategies are of the threshold type - the players quit when the corresponding rates of arrival reach a certain threshold from above;

in addition, there may exist a time interval such that (q_1^t, q_2^t) continuously decrease from one to zero on this interval.

We show that if $R_2 \geq R_3$, then, when a certain critical level of $\lambda_d(\bar{\pi}, t)$ is reached, it is optimal for both players to exit. If $R_2 < R_3$, then the simultaneous exit is not an equilibrium, and three types of equilibria are possible.

Assume that if the project is "good" and both players keep experimenting, the players are equally likely to observe the first success. Let $G_i(\bar{\pi}, t) = 0.5\lambda_d(\bar{\pi}, t)(R_1 + R_2) - rC$ denote the instantaneous expected payoff flow of player i if none of the players stopped until time $t > 0$. Let $F_i(\bar{\pi}, t)$ denote the expected value of player i if player j stopped at time t, and player i did not. From now on, we will consider the simple strategies q_1, q_2, whose singular continuous components of the Lebesgue decompositions are trivial. Given the strategy profile (q_i, q_j), the value of player i in the game that starts at $t = 0$ is

$$\begin{aligned} V_i(\bar{\pi}; q_i, q_j) &= \int_0^\infty e^{-rt} p_d(\bar{\pi}, t) G_i(\bar{\pi}, t) q_i(t) q_j(t) dt \\ &+ \int_{\{t \geq 0 \ | \ \Delta q_i(t) = 0\}} e^{-rt} p_d(\bar{\pi}, t) F_i(\bar{\pi}, t) q_i(t)(-dq_j(t)) \\ &+ \int_0^\infty e^{-rt} p_d(\bar{\pi}, t) R_3 q_j(t)(-dq_i(t)). \end{aligned} \tag{5.2}$$

Later we will show that $F_i(\bar{\pi}, \cdot)$ is continuous and has the finite limit as $t \to \infty$, hence, $V(\bar{\pi}; q_i, q_j)$ is well-defined and finite. Note that the second integral in (5.2) takes into account jumps in q_j only, and last integral takes into account jumps in q_i only as well as simultaneous jumps in q_i and q_j. We will use the same equilibrium concepts as in Definition 2.6.

Once one of the players has quitted experimentation, the other player faces a non-strategic stopping problem, which can be easily solved. Thus, when considering subgame perfect equilibria, we will first examine subgames when one of the players has stopped, and then move to subgames where neither player has quitted as yet. To simplify the notation, we suppress the dependence of value functions on the other player's strategy. Since the players are symmetric, we also drop the subscripts identifying the players.

5.3. **Follower's problem.** Consider a subgame that starts after the history such that only one of the players has stopped. Suppose, this happened at time t. Then the remaining player (the follower) chooses a strategy q_f^t satisfying the conditions of Definition 2.2, which solves the following problem:

$$\begin{aligned} F(\bar{\pi}, t) &= \sup_{q_f^t} \left[\int_t^\infty e^{-r(t'-t)} q_f^t(t') \frac{p_s(\pi_s(\bar{\pi}, t), t')}{p_d(\bar{\pi}, t)} \left(\lambda_s(\pi_s(\bar{\pi}, t), t') R_1 - rC \right) dt' \right. \\ &\left. + R_3 \int_t^\infty e^{-r(t'-t)} \frac{p_s(\pi_s(\bar{\pi}, t), t')}{p_d(\bar{\pi}, t')} (-dq_f^t(t')) \right], \end{aligned} \tag{5.3}$$

where the first integral is the expected present value (EPV) of the payoff while the follower is active, and the second integral is the EPV of the payoff when the follower exits prior to the first failure.

Introduce the notation

$$A = \frac{r(C+R_3)}{R_1-R_3}.$$

Similarly to Lemma 3.1, we obtain

Lemma 5.1. *The value of the follower, given by equation* (5.3), *can be equivalently written as*

$$\begin{aligned} F(\bar\pi,t) &= R_3 \qquad\qquad (5.4)\\ &+\frac{(R_1-R_3)e^{rt}}{p_d(\bar\pi,t)}\sup_{q_f^t}\int_t^\infty e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\left(\lambda_s(\pi_s(\bar\pi,t),t')-A\right)q_f^t(t')dt'. \end{aligned}$$

The proof is analogous to the proof of Lemma 3.1. The first term in representation (5.1) is the value of immediate exit; the second term is the option value of waiting. Since $\lambda_s(\pi_s(\bar\pi,t),t')\to 0$ as $t'\to+\infty$, the follower exits either instantly if either $\lambda_s(\pi_s(\bar\pi,t),t')\le A$ and t is to the right from the point $\hat t_s = \arg\max_{t'\ge t}\lambda_s(\pi_s(\bar\pi,t),t')$; or $t<\hat t_s$ and $\max_{T\ge t}\Phi(A,\bar\pi,t;T)\le 0$, where

$$\Phi(A,\bar\pi,t;T) = (R_1-R_3)\int_t^T e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\left(\lambda_s(\pi_s(\bar\pi,t),t')-A\right)dt';$$

(in the case of equality, it is also optimal to wait until the local maximizer T is achieved). Otherwise, the follower exits at time $T_s = T_s(A,\pi_s(\bar\pi,t)) < +\infty$, $T_s>t$, which is the largest solution of the following equation

$$\lambda_s(\pi_s(\bar\pi,t),T) = A. \tag{5.5}$$

Thus, the follower's exit time $t_f = t_f(A,\pi_s(\bar\pi,t))$ can equal to t or $T_s(A,\pi_s(\bar\pi,t))$. In a generic case, we can write the follower's optimal strategy as

$$\hat q_f^t(t') = \begin{cases} 1, & \forall\, t\le t'\le t_f(A,\pi_s(\bar\pi,t)),\\ 0, & \forall\, t'>t_f; \end{cases}$$

and the follower's value as

$$F(\bar\pi,t) = S + \frac{e^{rt}}{p_d(\bar\pi,t)}\Phi(A,\bar\pi,t;t_f(A,\pi_s(\bar\pi,t))). \tag{5.6}$$

5.4. **Value of player** i**.** Consider a subgame starting at $t\ge 0$ after a history such that none of the players has yet acted. Consider the value function of player i in such a subgame. We have

$$
\begin{aligned}
V_i(\bar{\pi},t;q_i^t,q_j^t) &= \int_t^\infty e^{-r(t'-t)}\frac{p_d(\bar{\pi},t')}{p_d(\bar{\pi},t)}G_i(\bar{\pi},t')q_i^t(t')q_j^t(t')dt' \\
&+ \int_{\{t'\geq t \ | \ \Delta q_i^t(t')=0\}}^\infty e^{-r(t'-t)}\frac{p_d(\bar{\pi},t')}{p_d(\bar{\pi},t)}F_i(\bar{\pi},t')q_i^t(t')(-dq_j^t(t')) \\
&+ \int_t^\infty e^{-r(t'-t)}\frac{p_d(\bar{\pi},t')}{p_d(\bar{\pi},t)}R_3 q_j^t(t')(-dq_i^t(t')),
\end{aligned} \tag{5.7}
$$

$F_i(\bar{\pi},t')$ is the value of the follower, and $G_i(\bar{\pi},t') = 0.5\lambda_d(\bar{\pi},t')(R_1+R_2) - rC$ is the expected payoff flow that player i gets when both players are experimenting. Introduce

$$
\kappa = 1 + \frac{R_2 - R_3}{R_1 - R_3}. \tag{5.8}
$$

If $R_2 \geq R_3$ (even if you know that you will be not the first to succeed, it is non-optimal to give up the project of unknown quality), then $\kappa \geq 1$. If $R_2 < R_3$ (it is better not to succeed if the project is of unknown quality, then when the project is known to be "good"), then $0 < \kappa < 1$. If $R_2 \uparrow R_1$, then $\kappa \uparrow 2$. (Recall that $R_1 = R_1$ means that no loss if you are not the first to succeed if the project is "good").

The following result can be proved in the same manner as Lemma 3.8.

Lemma 5.2. *We have*

$$
\begin{aligned}
V_i(\bar{\pi},t;q_i^t,q_j^t) &= R_3 \\
&+ \frac{e^{rt}}{p_d(\bar{\pi},t)}\int_t^\infty \Big[(R_1-R_3)e^{-rt'}p_d(\bar{\pi},t')\left(0.5\kappa\lambda_d(\bar{\pi},t')-A\right)q_j^t(t')\big] \\
&\qquad -\Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t'))(q_j^t(t'))'\Big]\, q_i^t(t')dt' \\
&+ \frac{e^{rt}}{p_d(\bar{\pi},t)} \sum_{\substack{t'\geq t:\\ \Delta q_i(t')=0\\ \Delta q_j(t')\neq 0}} \Phi(A,\bar{\pi},t';t_f(A,\bar{\pi},t'))q_i^t(t')(-\Delta q_j(t')) \\
&- \frac{e^{rt}}{p_d(\bar{\pi},t)} \sum_{\substack{t'\geq t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} R_3 e^{-rt'}\left(q_i^t(t')+\Delta q_i^t(t')\right)(-\Delta q_j(t')).
\end{aligned} \tag{5.9}
$$

Since $\lambda_d(\bar{\pi},t') \to 0$ as $t' \to +\infty$, it is non-optimal to experiment jointly either after time t or after time $T = t_d(\kappa,A,\bar{\pi}) < +\infty$, which is the largest solution of the following equation

$$
0.5\kappa\lambda_d(\bar{\pi},T) = A. \tag{5.10}
$$

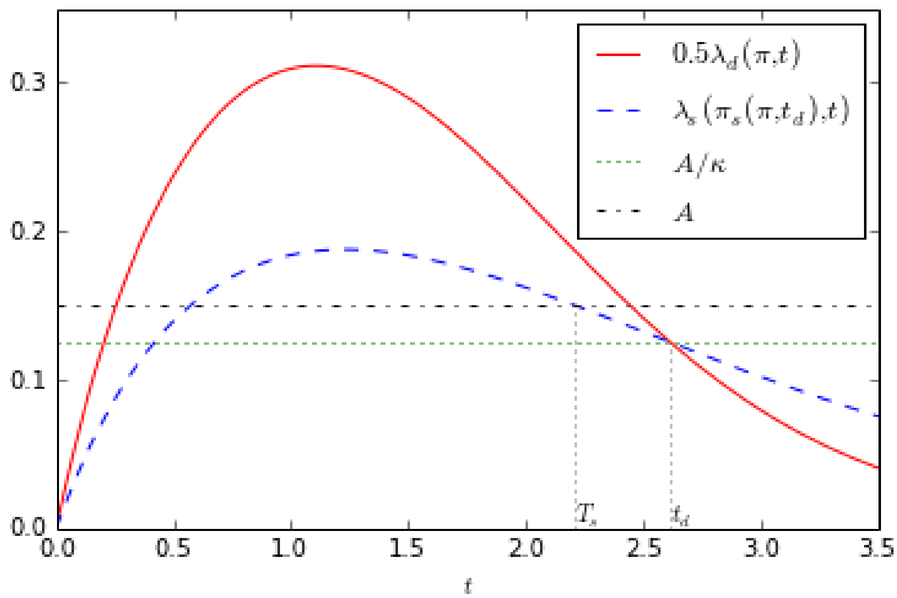

FIGURE 6. Illustration for Lemma 5.3 (c). Parameters: $\lambda = 1$, $A = 0.15$ $\bar{\pi} = 0.25$ $\kappa = 1.2$.

The conditions for the instantaneous exit can be derived in the same way as in the case of failures. We consider the more interesting case, when $t < t_d(\kappa, A, \bar{\pi})$, and, for all $t < t_d(\kappa, A, \bar{\pi})$, it is non-optimal to stop the joint experimentation.

Lemma 5.3. *Let $t < t_d(\kappa, A, \bar{\pi})$. Let $T_s(A, \pi_s(\bar{\pi}, t_d))$ denote the largest solution to (5.5) for $t = t_d(\kappa, A, \bar{\pi})$. Then*

(a) The graphs of functions $0.5\lambda_d(\bar{\pi}, t)$ and $\lambda_s(\pi_s(\bar{\pi}, t_d), t)$ have two points of intersection: at $t = 0$ and $t = t_d(\kappa, A, \bar{\pi})$.

(b) For $t \in (0, t_d(\kappa, A, \bar{\pi}))$,

$$\lambda_s(\pi_s(\bar{\pi}, t_d), t) < 0.5\lambda_d(\bar{\pi}, t).$$

(c) If $\kappa \geq 1$, then $T_s(A, \pi_s(\bar{\pi}, t_d)) \leq t_d(\kappa, A, \bar{\pi})$, hence both players exit at $t_d(\kappa, A, \bar{\pi})$.

(d) If $\kappa < 1$, the $T_s(A, \pi_s(\bar{\pi}, t_d)) > t_d(\kappa, A, \bar{\pi})$, hence the leader exits at $t_d(\kappa, A, \bar{\pi})$, and the follower exits at $T_s(A, \pi_s(\bar{\pi}, t_d))$.

The proof of (a) and (b) is the same as the proof of Lemma 8.3. The validity of (c) and (d) follows from (a) and (b), and it is evident from Fig. 6 and Fig. 7, respectively.

Theorem 5.4. *Let $t < t_d = t_d(\kappa, A, \bar{\pi})$ and $R_2 \geq R_3$. Then there exist a unique SPE defined by*

$$\hat{q}_i^t(t') = \hat{q}_j^t(t') = \begin{cases} 1, & t' \in (t, t_d], \\ 0, & t' > t_d. \end{cases} \tag{5.11}$$

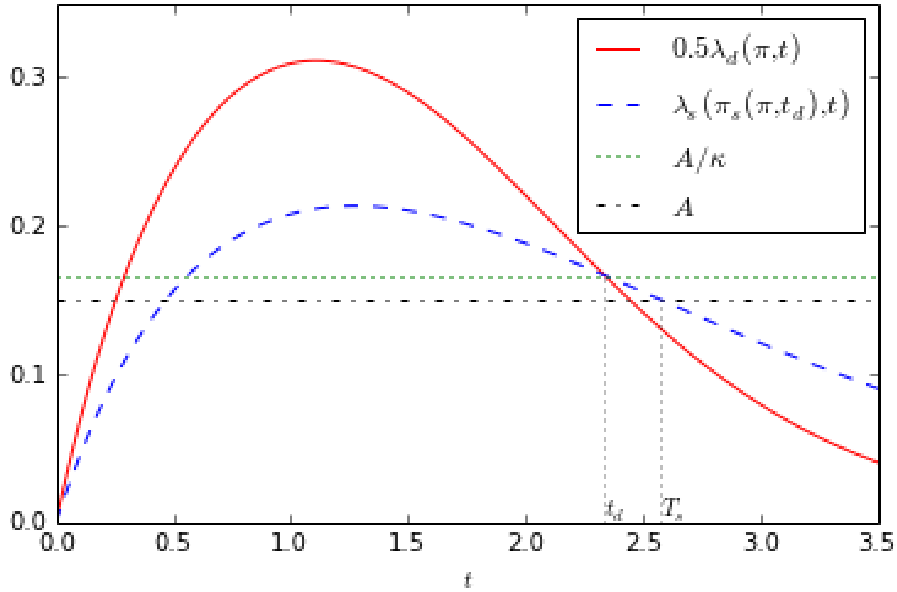

FIGURE 7. Illustration for Lemma 5.3 (d). Parameters: $\lambda = 1$, $A = 0.15$ $\bar{\pi} = 0.25$ $\kappa = 0.9$.

The players' payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = R_3 + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d), \tag{5.12}$$

where

$$\Psi(\kappa, A, \bar{\pi}, t; T) = \int_t^T (R_1 - R_3) e^{-rt'} p_d(\bar{\pi}, t') \left(0.5\kappa\lambda_d(\bar{\pi}, t') - A\right) dt'. \tag{5.13}$$

Let $R_2 < R_3$. As in the case of failures, let $\hat{T}_f(A, \bar{\pi}) = \inf\{t' \geq t_d \mid t_f(A, \pi_s(\bar{\pi}, t')) = t'\}$. Then symmetric equilibria are formulated in terms of the solution of the Cauchy problem.

$$\begin{aligned} U_0(t') &:= U_0(A, \bar{\pi}, t') = (R_1 - R_3) e^{-rt'} p_d(\bar{\pi}, t') \left(\kappa\lambda_d(\bar{\pi}, t') - A\right), \\ U_1(t') &:= \Psi(\kappa, A, \bar{\pi}, t', t_f(A, \pi_s(\bar{\pi}, t'))), \end{aligned}$$

and consider the Cauchy problem (4.8) subject to $\alpha(t, t) = 1$. The statement and proof of Lemma 4.8 are repeated word by word.

Theorem 5.5. *Let $R_2 < R_3$. Consider a subgame that starts at $t \in [t_d(\kappa, A, \bar{\pi}), \hat{T}_f(A, \bar{\pi}; t))$ after a history such that none of the players has yet acted. Then, there exists a symmetric SPE given by the following pair of simple consistent strategies:*

$$\hat{q}_j^t(t') = \hat{q}_j^t(t') = \begin{cases} \alpha(t, t'), & t' \in (t, \hat{T}_f(A, \bar{\pi})) \\ 0, & t' \geq \hat{T}_f(A, \bar{\pi}). \end{cases}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = R_3$$

Proof is the same as the proof of Theorem 4.9.

Theorem 5.6. *Let $R_2 < R_3$. Then, in a subgame that starts at $0 \leq t < t_d = t_d(\kappa, A, \bar{\pi})$ after a history such that none of the players has yet acted, there exists a symmetric equilibrium:*

$$\hat{q}_j^t(t') = \hat{q}_j^t(t') = \begin{cases} 1, \ t' \in (t, t_d] \\ \alpha(t_d, t'), \ t' \in (t_d, \hat{T}_f(A, \bar{\pi})) \\ 0, \ t' \geq \hat{T}_f(A, \bar{\pi}). \end{cases} \tag{5.14}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = R_3 + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d). \tag{5.15}$$

The proof is the same as the proof of Theorem 4.10.

There are also asymmetric equilibria, characterized in the following Theorems, which can be proved in the same manner as Theorems 4.12 and 4.15.

Theorem 5.7. *Let $R_2 < R_3$. Then, in a subgame that starts at $t \in [t_d(\kappa, A, \bar{\pi}), \hat{T}_f(A, \bar{\pi}))$ after a history such that none of the players has yet acted, there are two asymmetric equilibria given by the following pairs of simple consistent strategies: for $(i, j) \in \{(1, 2), (2, 1)\}$,*

$$\hat{q}_i^t(t') = 0 \ \forall t' > t_d, \tag{5.16}$$

and

$$\hat{q}_j^t(t') = \begin{cases} 1, & t' \in (t, t_f(t_d)], \\ 0, & t' > t_f(t_d). \end{cases} \tag{5.17}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = R_3, \quad V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = F(\bar{\pi}, t).$$

Theorem 5.8. *Let $R_2 < R_3$. Then, in a subgame that starts at $0 \leq t < t_d = t_d(\kappa, A, \bar{\pi})$ after a history such that none of the players has yet acted, there are two* *asymmetric equilibria given by the following pairs of simple consistent strategies: for* $(i,j) \in \{(1,2),(2,1)\}$,

$$\hat{q}_i^t(t') = \begin{cases} 1, & t' \in (t, t_d], \\ 0, & t' > t_d, \end{cases} \tag{5.18}$$

and

$$\hat{q}_j^t(t') = \begin{cases} 1, & t' \in (t, t_f(t_d)], \\ 0, & t' > t_f(t_d). \end{cases} \tag{5.19}$$

The players payoffs are

$$V_i(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = R_3 + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d), \tag{5.20}$$

$$V_j(\bar{\pi}, t; \hat{q}_i^t, \hat{q}_j^t) = F(\bar{\pi}, t_d) + \frac{e^{rt}}{p_d(\bar{\pi}, t)} \Psi(\kappa, A, \bar{\pi}, t; t_d). \tag{5.21}$$

6. Extensions and generalization

6.1. **Hump-shaped distributions.** The results of the paper derived for the case when random times of news arrival are Erlang(2,λ) random variables can be generalized if we take as a primitive of the model the rate of arrival of a time-inhomogeneous Poisson process. Let the true rate of arrival of the news be $\lambda \in \{0, \lambda_1(t)\}$, where $\lambda_1(t)$ is the rate of arrival of some time inhomogeneous Poisson process. Common prior assigns probability $\bar{\pi}$ to $\lambda = 0$. Let τ be a random time of the news arrival and

$$p_{\lambda_1}(t) = prob(\tau > t) = e^{-\int_0^t \lambda_1(t')dt'}.$$

The expected arrival rate is

$$\lambda(\bar{\pi}; t) = \frac{(1-\bar{\pi})\lambda_1(t) p_{\lambda_1}(t)}{\bar{\pi} + (1-\bar{\pi}) p_{\lambda_1}(t)}.$$

Lemma 6.1. *a) Let* $\bar{\pi} \in (0,1)$ *and*

$$(1/\lambda_1)''(t) > 0, \quad \forall\, t > 0, \tag{6.1}$$

$$-(1/\lambda_1)'(+0) > \bar{\pi}. \tag{6.2}$$

Then

(i) *as a function of* t, $\lambda(\bar{\pi}, t)$ *has the only global maximum* $\hat{t} = \hat{t}(\bar{\pi})$;
(ii) $\lambda(\bar{\pi}, t)$ *is strictly increasing on* $[0, \hat{t}(\bar{\pi})]$ *and decreasing on* $[\hat{t}(\bar{\pi}), +\infty)$.
b) If (6.1) *holds but* (6.2) *fails, then* $\lambda(\bar{\pi}, t)$ *is a decreasing function on* $\mathbb{R}_+$.

Suppose that n players experiment with projects of unknown quality, and the rates of arrival of news are i.i.d. random variables. Let $\{\tau_i\}_{i=1}^n$ be i.i.d. random times of news arrival, and

$$p_{n\lambda_1} = \Pi_{i=1}^n prob(\tau_i > t) = p_{\lambda_1}^n.$$

The expected arrival rate is

$$\lambda_n(\bar\pi, t) = \frac{(1-\bar\pi)n\lambda_1(t)p_{n\lambda_1}(t)}{\bar\pi + (1-\bar\pi)p_{n\lambda_1}(t)}.$$

Lemma 6.2. *a) Let $\bar\pi \in (0,1)$ and*

$$(1/\lambda_1)''(t) \;>\; 0, \quad t > 0, \tag{6.3}$$
$$-(1/\lambda_1)'(+0) \;>\; n\bar\pi. \tag{6.4}$$

Then

(i) as a function of t, $\lambda_n(\bar\pi, t)$ has the only global maximum $\hat t_n = \hat t_n(\bar\pi)$;
(ii) $\lambda_n(\bar\pi, t)$ is strictly increasing on $[0, \hat t_n(\bar\pi)]$ and decreasing on $[\hat t_n(\bar\pi), +\infty)$.
b) If (6.3) holds but (6.4) fails, then $\lambda_n(\bar\pi, t)$ is a decreasing function on $\mathbb{R}_+$.

If $\bar\pi < -(1/\lambda_1)'(+0) < +\infty$, then the qualitative behavior of the arrival rate $\lambda_n(\bar\pi, t)$ depends on n.

For a sufficiently large n, the arrival rate $\lambda_n(\bar\pi, t)$ is a decreasing function, and the model is qualitatively the same as the exponential bandit model.

6.2. **Classification of one-humped bandits.** Next, we characterize different possible types of one-humped bandits. We start with the following preliminary remarks.

- If $-(1/\lambda_1)'(+0) = +\infty$, then (6.4) holds for any n.
- If $0 < -(1/\lambda_1)'(+0) < +\infty$, then $1/\lambda_1(t)$ is bounded as $t \to 0$, and $\lambda_1(0) > 0$.
- It is possible that $\lambda_1(0) > 0$, (6.3) holds but $-(1/\lambda_1)'(+0) = +\infty$.
- If $\lambda_1'(0)$ exists and (6.3) holds, then $\lambda_1(0) > 0 \Leftrightarrow 0 < -(1/\lambda_1)'(+0) < +\infty$.

Definition 6.3. Let $(1/\lambda_1)''(t) > 0, \quad \forall\, t > 0$.

We call the bandit model defined by λ_1 a one-humped model of Type I, II and III if the corresponding condition below holds

I. $\lambda_1(0) = 0$, and $\lambda_1'(0)$ exists, and it is finite;
II. $\lambda_1(0) > 0$ and $\lambda_1'(0)$ exists, and it is finite;
III. $\lambda_1(0) > 0$ and $-(1/\lambda_1)'(+0) = +\infty$.

6.3. **One-humped bandits of Type I: further properties.** An example: Erlang-k bandits, $k \ge 2$.

Type I bandits have the properties that we used to study Erlang-2 bandits:

1. $\lambda_n(\bar\pi, t)/n < \lambda(\bar\pi, t), \ \forall\, t > 0, n > 1$.
2. For any $\kappa > 1, n > 1$, there exists $t(\kappa, n) > 0$ such that

$$\kappa\lambda_n(\bar\pi, t)/n > \lambda(\bar\pi, t), \quad 0 < t < t(\kappa, n).$$

3. Let $\kappa \in (1, n)$, and let $\tilde t = \tilde t(\kappa, n) > 0$ be a solution of the equation

$$\kappa\lambda_n(\bar\pi, t)/n = \lambda(\bar\pi, t).$$

Then $\tilde t(\kappa, n)$ exists, and it is unique.

6.4. **One-humped bandits of Types II and III.** Properties listed for Type I hold, and equilibria of the same types are possible.

Depending on the parameters, the usual encouragement effect can be observed (as in exponential bandit models).

An additional effect and type of equilibria (if r is sufficiently large): discouragement (crowding out) effect: $\exists m \geq 1$ s.t.

(1) if $n < m$ players are in the game at time 0, they will find it optimal to start experimenting with the "bad news" technology;
(2) $n \geq m$ players will not start experimenting unless $n - m$ of them exit instantly.

In the game starting at $t = 0$, mixed equilibria similar to the ones considered in a subgame that starts at the optimal leader threshold (in the model with breakdowns) are possible.

6.5. **Multi-humped bandits.** We call the model a multi-humped model, if λ has more than one point of local maximum. Examples

(a) the environment with some seasonality;
(b) if business cycle effects are taken into account;
(c) endogenous multi-humped bandits.

6.6. **Endogenous multi-humped bandits.** Assume that the players plan to enter the game with breakdowns at times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n < t_{n+1} := +\infty$; this can be an equilibrium outcome if, for example, players are asymmetric.

Then the rate of arrival $\Lambda(\bar{\pi}; t)$, which the players that are in the game face, is defined as follows. For $k = 1, 2, \ldots, n$ and $t \in [t_k, t_{k+1})$,

$$\Lambda(\bar{\pi}; t) = \sum_{j=1}^{k} \lambda(\bar{\pi}; t - t_j).$$

Clearly, more than one hump is possible, and if the underlying one-humped bandit is of Type II or III, then Λ exhibits jumps. Furthermore, in the model with breakdowns, the expected cost C_t at time t is not of the form $C\kappa_{n-1}\lambda_n(\bar{\pi}, t)/(n)$ ($\kappa_{n-1} := 1 + (n-1)L/C$). For player m, $m \leq n$, the expected cost is

$$C_{m,t} = C\lambda(\bar{\pi}; t - t_m) + L \sum_{1 \leq j \leq n, j \neq m} \lambda(\bar{\pi}; t - t_j).$$

7. Conclusion

I suggested a new model for strategic experimentation, where "good" or "bad" news arrive at random times which are modeled as i.i.d. Erlang(2,λ) variables. The initial value of the parameter λ is not known, and can be either positive or zero. Erlang bandits models are almost as tractable as exponential bandits models and can incorporate such realistic features as dependence of the expected rate of news arrival on the time elapsed since the start of an experiment and judgement about the quality of a "risky" arm based on evidence of a series of trials as opposed to a single evidence of success or failure as in exponential models with conclusive experiments.

I considered strategic experimentation with Erlang bandits and

- Characterized SPE in a model with conclusive failures.
- Characterized optimal stopping strategies in terms of critical levels of expected rates of arrival of time-inhomogeneous Poisson processes.
- Showed that depending on parameters of the model the following equilibrium outcomes are possible:
  (i) none of the players find it optimal to stop unless the first failure happens,
  (ii) the leader stops in finite time, and the follower either exits later in finite time, or never unless the first failure happens;
  (iii) a symmetric equilibrium, where the players do not stop until the optimal stopping time of the leader and then randomize in the interval between the leader's and the follower's optimal stopping times;
  (iv) players stop simultaneously (special case).
- Similar results hold in a model with breakthroughs.
- In a model with costly failures, it may be better to "become an ostrich" in the sense that a player, who has to stop first in the asymmetric SPE, is better off experimenting alone.
- Suggested a classification of humped bandits
- Showed that a wide class of humped bandits enjoy the same properties as Erlang bandits.

In the future, I plan to extend the Erlang bandits model to inconclusive experiments, correlated arms, private payoffs, and other types of humped bandits.

References

[1] D. Bergemann, and U. Hege, "Dynamic venture capital financing, learning and moral hazard," *Journal of Banking and Finance*, 22 (1998), 703-735.
[2] D. Bergemann, and U. Hege, "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36:4 (2005), 719-752.
[3] D. Bergemann, and J. Välimäki, "Learning and strategic pricing," *Econometrica*, 64 (1996), 1125-1149.
[4] D. Bergemann, and J. Välimäki, "Experimentation in markets," *Review of Economics Studies*, 67 (2000), 213-234.
[5] D. Bergemann, and J. Välimäki, "Dynamic price competition," *Journal of Economic Theory*, 127 (2006), 232-263.
[6] D. Bergemann, and J. Välimäki, "Bandit problems." In The New Palgrave Dictionary of Economics (Steven N. Durlauf and Larry E. Blume, eds.), PalgraveMacmillan, Basingstoke (2008).
[7] P. Bolton, and C. Harris, "Strategic experimentation", *Econometrica* 67 (1999), 349-374.
[8] A. Cohen, and E. Solan, "Bandit Problems with Lévy Payoff Processes," Mathematics of Operations Research, 38 (2013), 92?107.

[9] D. Cummins, and P.M. Barrieu, “Innovations in insurance markets: hybrid and securitized risk.” In Handbook of Insurance (Georges Dionne, ed.) Springer, 2nd ed. (2013)

[10] J.-P. Decamps, and T. Mariotti, “Investment timing and learning externalities”, *Journal of Economic Theory* 118 (2004), 80-102.

[11] J.-P. Decamps, T. Mariotti, and S.Villeneuve, “Investment timing under incomplete information”, *Mathematics of Operations Research* 30 (2005), 472-500.

[12] J.-P. Decamps, T. Mariotti, and S.Villeneuve, “Investment timing under incomplete information: Erratum”, *Mathematics of Operations Research* 34 (2009), 255-256.

[13] P.K. Dutta, and A. Rustichini “A Theory of Stopping Time Games With Applications to Product Innovations and Asset Sales, *Economic Theory*, 3 (1993), 743-763.

[14] P. Heidhues, S. Rady, and P. Strack, “Strategic experimentation with private payoffs,” *Journal of Economic Theory*, 159 (2015), 531-551.

[15] J. Hörner, and A. Skrzypacz, “Learning, experimentation and information design,” Working Paper, 2016.

[16] T.D. Gerarden, R.G. Newell, R.N. Stavins, and R.C. Stowe, “An assessment of the energy-efficiency gap and its implications for climate change policy,” NBER Working Paper 20905 http://www.nber.org/papers/w20905

[17] J. Hörner, N.A. Klein, and S. Rady, “Strongly symmetric equilibria in bandit games,” *Cowles Foundation Discussion Paper No. 1056* (2014), Available at SSRN: http://ssrn.com/abstract=2482335

[18] G. Keller, S. Rady, and M. Cripps, “Strategic Experimentation with Exponential Bandits,” *Econometrica*, 73:1 (2005), 39-68.

[19] G. Keller, and S. Rady, “Optimal experimentation in a changing environment,” *Review of Economic Studies*, 66 (1999), 475-507.

[20] G. Keller, and S. Rady, “Strategic Experimentation with Poisson Bandits,” *Theoretical Economics*, 5 (2010), 275-311.

[21] G. Keller, and S. Rady, “Breakdowns,” *Theoretical Economics*, 10 (2015), 175-202.

[22] N. Klein, and S. Rady, “Negatively correlated bandits,” *Review of Economic Studies*, 78:2 (2008), 693-732.

[23] U. Khan, and M.B. Stinchcombe, “The Virtues of Hesitation: Optimal Timing in a Non-stationary World,” *American Economic Review*, 105:3 (2015), 1147-1176.

[24] R. Laraki, E. Solan, and N. Vieille, “Continuous-time Games of Timing,” *Journal of Economic Theory*, 120 (2005), 206–238.

[25] M. Ludkovski, and R. Sircar, “Exploration and Exhaustibility in Dynamic Cournot Games,” *European Journal of Applied Mathematics*, 23 (2012), 343-372.

[26] P. Murto, and J. Välimäki, “Delay and information aggregation in stopping games with private information,” *Journal of Economic Theory*, 148 (2013), 2404-2435.

[27] K. Roberts, and M.L. Weitzman, “Funding criteria for research, development, and exploration projects,” *Econometrica*, 49:5 (1981), 1261-1288.

[28] D. Rosenberg, A. Salomon, and N. Vieille, “On games of strategic experimentation,” *Games and Economic Behavior*, 82 (2013), 31-51.

[29] M. Rothschild, “A two-armed bandit theory of market pricing,” *Journal of Economic Theory*, 9:2 (1974), 185-202.

[30] M.L. Weitzman, “Optimal search for the best alternative,” *Econometrica*, 47:3 (1979), 641-654.

8. APPENDIX

8.1. **Proof of Lemma 3.1.** We start with the following result.

Lemma 8.1. *Let the following conditions hold:*

(i) $f : [0, +\infty) \to \mathbb{R}$ *be continuous and differentiable;*
(ii) there exists $B > 0$ *and* $r > 0$ *s.t.* $|f(t)| \le Be^{-rt}\ \forall\, t \ge 0$;
(iii) q *satisfies the condtions of Definition 2.1 and the singular continuous component of the Lebesgue decomposition of* q *is trivial.*

Then

$$\int_t^\infty f(t')(-dq(t')) = f(t)q(t) + \int_t^\infty f'(t')q(t')dt'. \tag{8.1}$$

Proof. By definition,

$$\int_t^\infty f(t')(-dq(t')) = -\int_t^\infty f(t')q'(t')dt' + \sum_{\substack{t' \ge t:\\ \Delta q(t') \ne 0}} f(t')(-\Delta q(t')),$$

and

$$\begin{aligned}
-f(t)q(t) &= \int_t^\infty d\left(f(t')q(t')\right) \\
&= \int_t^\infty \left(f(t')q(t')\right)' dt' - \sum_{\substack{t' \ge t:\\ \Delta q(t') \ne 0}} f(t')(-\Delta q(t')) \\
&= \int_t^\infty f'(t')q(t')dt' + \int_t^\infty f(t')q'(t')dt' - \sum_{\substack{t' \ge t:\\ \Delta q(t') \ne 0}} f(t')(-\Delta q(t')) \\
&= \int_t^\infty f'(t')q(t')dt' - \int_t^\infty f(t')(-dq(t')).
\end{aligned}$$

Equation (8.1) follows. □

Using Lemma 8.1, we rewrite the second integral in (3.1) as

$$
\begin{aligned}
& S\int_t^\infty e^{-r(t'-t)}\frac{p_s(\pi_s(\bar\pi,t),t')}{p_d(\bar\pi,t)}(-dq^t(t')) \\
&= -S\frac{e^{rt}}{p_d(\bar\pi,t)}e^{-rt'}p_s(\pi_s(\bar\pi,t),t')q^t(t')\Big|_t^\infty + S\frac{e^{rt}}{p_d(\bar\pi,t)}\int_t^\infty \frac{d}{dt'}\left(e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\right)q^t(t')dt' \\
&= S + S\frac{e^{rt}}{p_d(\bar\pi,t)}\int_t^\infty e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\left(-r+\frac{p_s'(\pi_s(\bar\pi,t),t')}{p_s(\pi_s(\bar\pi,t),t')}\right)q^t(t')dt' \\
&= S + S\frac{e^{rt}}{p_d(\bar\pi,t)}\int_t^\infty e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\left(-r-\lambda_s(\pi_s(\bar\pi,t),t')\right)q^t(t')dt'.
\end{aligned}
$$

Substitute the last expression for the the second integral in (3.1), then

$$
F(\bar\pi,t) = S + \frac{e^{rt}}{p_d(\bar\pi,t)}\sup_{q^t}\int_t^\infty e^{-rt'}p_s(\pi_s(\bar\pi,t),t')\left(r(R-S)-\lambda_s(\pi_s(\bar\pi,t),t')C\right)q^t(t')dt'.
$$

Dividing and multiplying the integrand by λC and using (3.2), we arrive at (3.1).

8.2. **Proof of Lemma 3.8.** We start with the following result.

Lemma 8.2. *Let the following conditions hold:*

(i) $f:[0,+\infty)\to\mathbb{R}$ *is continuous and differentiable;*
(ii) there exists $B>0$ *and* $r>0$ *s.t.* $|f(t)|\le Be^{-rt}\ \forall\, t\ge 0$;
(iii) q_i, q_j *satisfy the conditions of of Definition 2.1, and the singular continuous components of the Lebesgue decomposition of* q_i, q_j *are trivial.*

Then

$$
\begin{aligned}
& \int_t^\infty f(t')q_j(t')(-dq_i(t')) \qquad (8.2)\\
= \ & -f(t)q_i(t)q_j(t) + \int_t^\infty f'(t')q_i(t')q_j(t')dt' \\
& -\int_t^\infty f(t')q_i(t')(-dq_j(t'))dt' - \sum_{\substack{t'\ge t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')).
\end{aligned}
$$

Proof. Applying Definition 2.5, we obtain

$$
\begin{aligned}
-f(t)q_i(t)q_j(t) &= \int_t^\infty d\left(f(t')q_i(t')q_j(t')\right) \\
&= \int_t^\infty \left(f(t')q_i(t')q_j(t')\right)' dt' - \sum_{\substack{t' \ge t: \\ \Delta(q_i(t')q_j(t')) \neq 0}} f(t')(-\Delta(q_i(t')q_j(t'))) \\
&= \int_t^\infty f'(t')q_i(t')q_j(t')dt' + \int_t^\infty f(t')q_i'(t')q_j(t')dt' + \int_t^\infty f(t')q_j'(t')q_i(t')dt' \\
&\quad - \sum_{\substack{t' \ge t: \\ \Delta q_i(t') \neq 0}} f(t')q_j(t')(-\Delta q_i(t')) - \sum_{\substack{t' \ge t: \\ \Delta q_j(t') \neq 0}} f(t')q_i(t')(-\Delta q_j(t')) \\
&\quad - \sum_{\substack{t' \ge t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')) \\
&= \int_t^\infty f'(t')q_i(t')q_j(t')dt' - \int_t^\infty f(t')q_j(t')(-dq_i(t')) \\
&\quad - \int_t^\infty f(t')q_i(t')(-dq_j(t'))dt' - \sum_{\substack{t' \ge t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} f(t')(-\Delta q_i(t')\Delta q_j(t')).
\end{aligned}
$$

Equation (8.2) follows. □

Now we can prove Lemma 3.8, in several steps.

Step 1. Rewrite the second integral on the RHS of (3.17) as

$$
\begin{aligned}
&\int_{\{t' \ge t \ | \ \Delta dq_i^t(t')=0\}}^\infty e^{-r(t'-t)} \frac{p_d(\bar\pi, t')}{p_d(\bar\pi, t)} F_i(\bar\pi, t') q_i^t(t')(-dq_j^t(t')) \\
&= \int_t^\infty e^{-r(t'-t)} \frac{p_d(\bar\pi, t')}{p_d(\bar\pi, t)} F_i(\bar\pi, t') q_i^t(t')(-dq_j^t(t')) \\
&- \sum_{\substack{t' \ge t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} e^{-r(t'-t)} \frac{p_d(\bar\pi, t')}{p_d(\bar\pi, t)} F_i(\bar\pi, t') q_i^t(t')(-\Delta q_j(t')).
\end{aligned}
\tag{8.3}
$$

Substitute for $F_i(\bar{\pi}, t')$ on the RHS of equation (4.13). Then (8.3) becomes

$$\begin{aligned}
&\int_{\{t' \geq t \ | \ \Delta dq_i^t(t')=0\}}^{\infty} e^{-r(t'-t)} \frac{p_d(\bar{\pi}, t')}{p_d(\bar{\pi}, t)} F_i(\bar{\pi}, t') q_i^t(t')(-dq_j^t(t')) \\
&= \frac{e^{rt}}{p_d(\bar{\pi}, t)} \int_t^{\infty} (Se^{-rt'} p_d(\bar{\pi}, t') + \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t'))) q_i^t(t')(-dq_j^t(t')) \qquad (8.4) \\
&- \frac{e^{rt}}{p_d(\bar{\pi}, t)} \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} (Se^{-rt'} p_d(\bar{\pi}, t') + \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t'))) q_i^t(t')(-\Delta q_j(t')).
\end{aligned}$$

By definition,

$$\begin{aligned}
&\int_t^{\infty} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t')) q_i^t(t')(-dq_j^t(t')) \\
&= \ - \int_t^{\infty} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t')) q_i^t(t')(q_j^t(t'))' dt' \\
&\quad + \sum_{\substack{t' \geq t: \\ \Delta q_j(t') \neq 0}} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t')) q_i^t(t')(-\Delta q_j(t')).
\end{aligned}$$

Therefore, we can rewrite (8.4) as

$$\begin{aligned}
&\int_{\{t' \geq t \ | \ \Delta dq_i^t(t')=0\}}^{\infty} e^{-r(t'-t)} \frac{p_d(\bar{\pi}, t')}{p_d(\bar{\pi}, t)} F_i(\bar{\pi}, t') q_i^t(t')(-dq_j^t(t')) \qquad (8.5) \\
= \ &\frac{e^{rt}}{p_d(\bar{\pi}, t)} \int_t^{\infty} (Se^{-rt'} p_d(\bar{\pi}, t') q_i^t(t')(-dq_j^t(t')) \\
&- \frac{e^{rt}}{p_d(\bar{\pi}, t)} \sum_{\substack{t' \geq t: \\ \Delta q_i(t') \neq 0 \\ \Delta q_j(t') \neq 0}} Se^{-rt'} p_d(\bar{\pi}, t') q_i^t(t')(-\Delta q_j(t')) \qquad (8.6) \\
&- \int_t^{\infty} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t'))) q_i^t(t')(q_j^t(t'))' dt' \\
&+ \frac{e^{rt}}{p_d(\bar{\pi}, t)} \sum_{\substack{t' \geq t: \\ \Delta q_i(t') = 0 \\ \Delta q_j(t') \neq 0}} \Phi(A, \bar{\pi}, t'; t_f(A, \bar{\pi}, t')) q_i^t(t')(-\Delta q_j(t')).
\end{aligned}$$

Step 2. Using Lemma 8.2 write the last integral in (3.17) as

$$\begin{aligned}
&\int_t^\infty e^{-r(t'-t)}\frac{p_d(\bar\pi,t')}{p_d(\bar\pi,t)}Sq_j^t(t')(-dq_i^t(t')) \qquad (8.7)\\
&=\frac{Se^{rt}}{p_d(\bar\pi,t)}\left[-e^{-rt'}p_d(\bar\pi,t')q_j^t(t')q_i^t(t')\Big|_t^\infty+\int_t^\infty\left(e^{-rt'}p_d(\bar\pi,t')\right)'q_i^t(t')q_j^t(t')dt'\right.\\
&\left.-\int_t^\infty e^{-rt'}p_d(\bar\pi,t')q_i^t(t')(-dq_j^t(t'))dt'-\sum_{\substack{t'\ge t:\\ \Delta q_i^t(t')\neq 0\\ \Delta q_j^t(t')\neq 0}}e^{-rt'}p_d(\bar\pi,t')(-\Delta q_i^t(t')\Delta q_j^t(t'))\right].
\end{aligned}$$

Using the fact that $q_i^t(t)=q_j^t(1)=1$, we rewrite (8.7) as

$$\begin{aligned}
&\int_t^\infty e^{-r(t'-t)}\frac{p_d(\bar\pi,t')}{p_d(\bar\pi,t)}Sq_j^t(t')(-dq_i^t(t'))=S \qquad (8.8)\\
&+\frac{Se^{rt}}{p_d(\bar\pi,t)}\left[\int_t^\infty e^{-rt'}p_d(\bar\pi,t')\left(-r-\lambda_d(\bar\pi,t')\right)q_i^t(t')q_j^t(t')dt'\right.\\
&\left.-\int_t^\infty e^{-rt'}p_d(\bar\pi,t')q_i^t(t')(-dq_j^t(t'))dt'-\sum_{\substack{t'\ge t:\\ \Delta q_i^t(t')\neq 0\\ \Delta q_j^t(t')\neq 0}}e^{-rt'}p_d(\bar\pi,t')(-\Delta q_i^t(t')\Delta q_j^t(t'))\right].
\end{aligned}$$

Adding (8.5) and (8.8) to the first integral in (3.17), we obtain

$$\begin{aligned}
V_i(\bar\pi,t;q_i^t,q_j^t) &= S+\frac{e^{rt}}{p_d(\bar\pi,t)}\\
&\times\int_t^\infty\left[e^{-rt'}p_d(\bar\pi,t')\left(r(R-S)-\frac{\kappa\lambda_d(\bar\pi,t')C}{2}\right)q_j^t(t')\right.\\
&\qquad\left.-\Phi(A,\bar\pi,t';t_f(A,\bar\pi,t'))(q_j^t(t'))'\right]q_i^t(t')dt'\\
&+\frac{e^{rt}}{p_d(\bar\pi,t)}\sum_{\substack{t'\ge t:\\ \Delta q_i(t')=0\\ \Delta q_j(t')\neq 0}}\Phi(A,\bar\pi,t';t_f(A,\bar\pi,t'))q_i^t(t')(-\Delta q_j(t'))\\
&-\frac{e^{rt}}{p_d(\bar\pi,t)}\sum_{\substack{t'\ge t:\\ \Delta q_i(t')\neq 0\\ \Delta q_j(t')\neq 0}}Se^{-rt'}\left(q_i^t(t')+\Delta q_i^t(t')\right)(-\Delta q_j(t')).
\end{aligned}$$

It remains to use the notation $A = r(R-C)/C$ to arrive at (3.18).

8.3. **Proof of Theorem 4.1.** Notice that the proposed strategies are consistent. Consider a subgame that starts at $t \geq 0$ after such a history that no player has acted as yet. Let player j follow the prescribed strategy (4.3). Then $\Delta q_j^t(t') = 0$ and $(q_j^t(t'))' = 0$ for all $0 \leq t \leq t'$. Player i chooses the best response $q_i^t(t')$ which solves the following problem

$$\sup_{q_i^t(t')} C \int_t^\infty e^{-rt'} p_d(\bar{\pi}, t') \left(A - 0.5\kappa\lambda_d(\bar{\pi}, t')\right) q_i^t(t') dt'. \tag{8.9}$$

Let $\psi(A, \bar{\pi}, t)$ be given by (4.6). Since $A \geq \hat{A}_d(\bar{\pi})$, $\psi(A, \bar{\pi}, t') \geq 0$ for all $t' \geq t$. Hence the best response of player i to q_j^t given by (4.3) is to play $\hat{q}^t(t') = 1$ for all $t' \geq t$. Hence $(\hat{q}_i^t, \hat{q}_j^t)$ given by (4.3) is a SPE.

8.4. **Proof of Lemma 4.2.** Straightforward calculations show that $\lambda_d'(\bar{\pi}, t) = 0$ iff t satisfies the following equation:

$$\bar{\pi}\frac{2(\lambda t)^2 - 1}{(\lambda t + 1)^2} = (1 - \bar{\pi})e^{-2\lambda t}. \tag{8.10}$$

The LHS in (8.10) is an increasing function, which is equal to $-\bar{\pi}$ at $t = 0$, and tends to $2\bar{\pi}$ as $t \to \infty$. The RHS is a decreasing exponential function, that equals $1 - \bar{\pi}$ at $t = 0$ and tends to zero as $t \to \infty$. Hence the equation (8.10) has a unique solution, denote it $\hat{t_d}$. Since the LHS in (8.10) is non-positive for $t \leq 1/(\lambda\sqrt{2})$, we conclude that $\hat{t_d} > 1/(\lambda\sqrt{2})$. Furthermore, $\lambda_d'(\bar{\pi}, t) > 0$ for all $t < \hat{t_d}$, and $\lambda_d'(\bar{\pi}, t) < 0$ for all $t > \hat{t_d}$, hence $\hat{t_d} = \hat{t_d}(\bar{\pi})$ is the global maximum of $\lambda_d(\bar{\pi}, t)$ on $[0, \infty)$.

Thus, $\lambda_d(\bar{\pi}, t)$ is increasing in t on $\{0 \leq t < \hat{t}_d(\bar{\pi})\}$, and decreasing in t on $\{t > \hat{t}_d(\bar{\pi})\}$. Hence for any $A < \hat{A}_d(\bar{\pi})$, equation (4.5) has two solutions. Denote them $t_d^* = t_d^*(\kappa, A, \bar{\pi}) < t_{*d} = t_{*d}(\kappa, A, \bar{\pi})$.

$$\frac{d\Psi(A, \bar{\pi}, t; T)}{dT} = Ce^{-rT} p_d(\bar{\pi}, T)\psi(\bar{\pi}, T).$$

If $T < t_d^*(\kappa, A, \bar{\pi})$ or $T > t_{*d}(\kappa, A, \bar{\pi})$, then $\psi(\bar{\pi}, T) > 0$, hence $\Psi(A, \bar{\pi}, t; T)$ is increasing for such T. If $t_d^*(\kappa, A, \bar{\pi}) < T < t_{*d}(\kappa, A, \bar{\pi})$, then $\psi(\bar{\pi}, T) < 0$, hence $\Psi(A, \bar{\pi}, t; T)$ is decreasing for such T. Hence t_d^* is the local maximum, and t_{*d} is the local minimum of $\Psi(A, \bar{\pi}, t; \dot{})$.

The part a) is proved. To prove b), we need to make a more detailed argument. Substitute (2.3) for $\lambda_d(\bar{\pi}, t)$ in (4.5) and write equation (4.5) as

$$\frac{A}{\lambda\kappa} = \frac{(1 - \bar{\pi})\lambda t(\lambda t + 1)}{\bar{\pi}e^{2\lambda t} + (1 - \bar{\pi})(\lambda t + 1)^2} \tag{8.11}$$

Denote $y := \lambda t$ and introduce

$$g(\bar{\pi}, y) := \frac{(1 - \bar{\pi})y(y + 1)}{\bar{\pi}e^{2y} + (1 - \bar{\pi})(y + 1)^2}. \tag{8.12}$$

Evidently, $g(\bar\pi, \lambda t)$ is the RHS in (8.11). Hence (8.11) is equivalent to

$$\frac{A}{\lambda\kappa} = g(\bar\pi, y). \tag{8.13}$$

Let $y_1 = y_1(\kappa, A, \bar\pi) := \lambda t^* d(\kappa, A, \bar\pi)$ be the smallest of the two solutions of (8.13).

Similarly, we can write (3.8) replacing $\lambda_s(\pi_s(\bar\pi, t^*_d)$ with (2.5):

$$\frac{A}{\lambda} = \frac{(1-\pi_s(\bar\pi, t^*_d))\lambda t}{\pi_s(\bar\pi, t^*_d)e^{\lambda t} + (1-\pi_s(\bar\pi, t^*_d))(\lambda t + 1)}. \tag{8.14}$$

Since

$$\pi_s(\bar\pi, t^*_d) = \frac{\bar\pi}{\bar\pi + (1-\bar\pi)(\lambda t^*_d + 1)e^{-\lambda t^*_d}} = \frac{\bar\pi}{\bar\pi + (1-\bar\pi)(y_1+1)e^{-y_1}},$$

we can write (8.14) in an equivalent way as

$$\frac{A}{\lambda} = f(\bar\pi, y_1(\kappa, A, \bar\pi), y), \tag{8.15}$$

where

$$f(\bar\pi, y_1(\kappa, A, \bar\pi), y) := \frac{(1-\bar\pi)y(y_1+1)}{\bar\pi e^{y_1+y} + (1-\bar\pi)(y_1+1)(y+1)}.$$

To finish the proof, we need the following lemma.

Lemma 8.3. *(i) Functions $g(\bar\pi, y)$ and $f(\bar\pi, y_1(\kappa, A, \bar\pi), y)$ intersect only at two points: $y = 0$, and $y = y_1$.*

(ii) $$g(\bar\pi, y) < f(\bar\pi, y_1(\kappa, A, \bar\pi), y) \;\Leftrightarrow\; y > y_1(\kappa, A, \bar\pi).$$

Proof. Write the equation $g(\bar\pi, y) = f(\bar\pi, y_1, y)$, where $y_1 = (\kappa, A, \bar\pi)$, using the definitions of f and g :

$$\frac{(1-\bar\pi)y(y+1)}{\bar\pi e^{2y} + (1-\bar\pi)(y+1)^2} = \frac{(1-\bar\pi)y(y_1+1)}{\bar\pi e^{y_1+y} + (1-\bar\pi)(y_1+1)(y+1)}. \tag{8.16}$$

Obviously, $y = 0$ is a solution of (8.16), and if $y > 0$, then (8.16) is equivalent to

$$\begin{aligned}
\frac{y+1}{\bar\pi e^{2y} + (1-\bar\pi)(y+1)^2} &= \frac{y_1+1}{\bar\pi e^{y_1+y} + (1-\bar\pi)(y_1+1)(y+1)} \\
&\Updownarrow \\
\bar\pi e^{y_1+y}(y+1) + (1-\bar\pi)(y_1+1)(y+1)^2 &= \bar\pi e^{2y}(y_1+1) + (1-\bar\pi)(y_1+1)(y+1)^2 \\
&\Updownarrow \\
e^{y_1}(y+1) &= e^{y}(y_1+1) \\
&\Updownarrow \\
\frac{e^y}{y+1} &= \frac{e^{y_1}}{y_1+1} \;\Leftrightarrow y = y_1.
\end{aligned}$$

Straightforward differentiation shows that

$$\begin{aligned}\frac{\partial g(\bar\pi,y)}{\partial y} &= \frac{1-\bar\pi}{(\bar\pi e^{2y}+(1-\bar\pi)(y+1)^2)^2}\left[(1-\bar\pi)(y+1)^2-\bar\pi e^{2y}(2y^2-1)\right]; \qquad (8.17)\\ \frac{\partial f(\bar\pi,y_1,y)}{\partial y} &= \frac{(1-\bar\pi)(y_1+1)}{(\bar\pi e^{y_1+y}+(1-\bar\pi)(y+1)(y_1+1))^2}\left[(1-\bar\pi)(y_1+1)-\bar\pi e^{y_1+y}(y-1)\right].\end{aligned}$$

We see that

$$\frac{\partial g(\bar\pi,0)}{\partial y} = 1-\bar\pi > \frac{(1-\bar\pi)(y_1+1)}{\bar\pi e^{y_1}+(1-\bar\pi)(y_1+1)} = \frac{\partial f(\bar\pi,y_1,0)}{\partial y}.$$

By continuity,

$$\frac{\partial g(\bar\pi,y)}{\partial y} > \frac{\partial f(\bar\pi,y_1,y)}{\partial y}$$

in a right neighborhood of zero. Hence $g(\bar\pi,y) > f(\bar\pi,y_1,y)$ in a right neighborhood of zero. Since there is only one positive intersection of these two functions, $g(\bar\pi,y) > f(\bar\pi,y_1,y)$ for all $y \in (0,y_1)$, and $g(\bar\pi,y) < f(\bar\pi,y_1,y)$ for all $y \in (y_1,\infty)$.

□

By Lemma 8.3,

$$f(\bar\pi,y_1,y_1) = g(\bar\pi,y_1) = \frac{A}{\lambda\kappa}.$$

If $\kappa = 1$, then y_1 is also the smallest solution of (8.15). If $\kappa > 1$ and (8.15) has two solutions, say, $z_1 < z_2$, then $z_1 > y_1$. Therefore, the followers optimal exit time is either z_1/λ or ∞, hence $t_f(A,\pi_s(\bar\pi,t_d^*)) > t_d^*$.

8.5. **Proof of Lemma 4.17.** Recall that if $A < A_d^*(\kappa,\bar\pi)$, the optimal stopping time of the leader is given by equation (8.13), where the function $g(\bar\pi,y)$ is given by (8.12). Since $A < A_d^*(\kappa,\bar\pi)$, the leader's optimal stopping time is finite, so if $t_s(A,\bar\pi) = +\infty$, then $t_s(A,\bar\pi) > t_d^*(A,\bar\pi)$. Consider the case when $t_s(A,\bar\pi) < +\infty$. Let

$$g_s(\bar\pi,y) = \frac{(1-\bar\pi)y}{\bar\pi e^y+(1-\bar\pi)(y+1)},$$

then equation (4.21) is equivalent to

$$\frac{A}{\lambda} = g_s(\bar\pi,y). \qquad (8.18)$$

Let $\kappa > 1$ and $\tilde z = \tilde z(\kappa,\bar\pi)$ be the positive root of the following quadratic equation:

$$\bar\pi z^2 - \kappa\bar\pi z - (\kappa-1)(1-\bar\pi) = 0. \qquad (8.19)$$

Let $\tilde y = \tilde y(\kappa,\bar\pi)$ be the only positive solution of the equation

$$\frac{e^{\tilde y}}{\tilde y} = \tilde z.$$

Lemma 8.4. *(i) Graphs of functions $\kappa g(\bar\pi,y)$ and $g_s(\bar\pi,y)$ intersect only at $y=0$, and $y=\tilde y$.*

$$(ii) \qquad \kappa g(\bar\pi,y)>g_s(\bar\pi,y) \ \Leftrightarrow\ y<\tilde y(\kappa,\bar\pi).$$

Proof. Write the equation $\kappa g(\bar\pi,y)=g_s(\bar\pi,y)$ using the definitions of g_s and g:

$$\frac{\kappa(1-\bar\pi)y(y+1)}{\bar\pi e^{2y}+(1-\bar\pi)(y+1)^2}=\frac{(1-\bar\pi)y}{\bar\pi e^{y}+(1-\bar\pi)(y+1)}. \tag{8.20}$$

Obviously, $y=0$ is a solution of (8.20). If $y>0$, then (8.20) is equivalent to

$$\begin{aligned}
\frac{\kappa(y+1)}{\bar\pi e^{2y}+(1-\bar\pi)(y+1)^2} &= \frac{1}{\bar\pi e^{y}+(1-\bar\pi)(y+1)}\\
&\Updownarrow\\
\kappa\bar\pi e^y(y+1)+(1-\bar\pi)\kappa(y+1)^2 &= \bar\pi e^{2y}+(1-\bar\pi)(y+1)^2\\
&\Updownarrow\\
\kappa\bar\pi\frac{e^y}{y+1}+(1-\bar\pi)\kappa &= \bar\pi\frac{e^{2y}}{(y+1)^2}+1-\bar\pi
\end{aligned}$$

Setting $z=e^y/(y+1)$, we arrive at equation (8.19), which has a single positive root. Straightforward differentiation shows that

$$\begin{aligned}
\kappa\frac{\partial g}{\partial y}(\bar\pi,y) &= \frac{\kappa(1-\bar\pi)}{(\bar\pi e^{2y}+(1-\bar\pi)(y+1)^2)^2}\left[(1-\bar\pi)(y+1)^2-\bar\pi e^{2y}(2y^2-1)\right];\\
\frac{\partial g_s}{\partial y}(\bar\pi,y) &= \frac{1-\bar\pi}{(\bar\pi e^{y}+(1-\bar\pi)(y+1))^2}\left[1-\bar\pi-\bar\pi e^y(y-1)\right].
\end{aligned}$$

We see that

$$\kappa\frac{\partial g}{\partial y}(\bar\pi,0)=\kappa(1-\bar p)>1-\bar\pi=\frac{dg_s}{dy}(\bar\pi,0).$$

By continuity,

$$\kappa\frac{\partial g}{\partial y}(\bar\pi,y)>\frac{\partial g_s}{\partial y}(\bar\pi,y)$$

in a right neighborhood of zero. Hence $\kappa g(\bar\pi,y)>g_s(\bar\pi,y)$ in a right neighborhood of zero. Since there is only one positive intersection of these two functions, $\kappa g(\bar\pi,y)>f(\bar\pi,y)$ for all $y\in(0,\tilde y)$, and $\kappa g(\bar\pi,y)<g_s(\bar\pi,y)$ for all $y\in(\tilde y,\infty)$.

□

It is easy to see that $\lambda_s(\bar\pi,t)$ is decreasing in $\bar\pi$ for any $t>0$. Since $\pi_s(\bar\pi,t)>\bar\pi$, we have $\lambda_s(\pi_s(\bar\pi,t),t')<\lambda_s(\bar\pi,t')$ for any $t'>0$. Hence, $f(\bar\pi,y_1,y)<g_s(\bar\pi,y)$ for any $y>0$ and any $y_1>0$. For $\kappa=1$, $f(\bar\pi,y_1,y_1)=g(\bar\pi,y_1)=A$, hence $g_s(\bar\pi,y_1)>A$ if $\kappa=1$, hence $y_s<y_1$.

Let

$$\hat y_d(\bar\pi)=\arg\max_{y\in\mathbb{R}_+}g(\bar\pi,y).$$

It is easy to see that if $g_s(\bar\pi, \hat y_d(\bar\pi)) < \kappa g(\bar\pi, \hat y_d(\bar\pi))$, then the point of intersection of the two functions $\tilde y > \hat y_d(\bar\pi)$, which immediately implies that $g_s(\bar\pi, y_1) < \kappa g(\bar\pi, y_1)$, because $y_1 < \hat y_d$, hence $y_1 < y_s$. Equivalently, $t^*_d < t_s$. Let

$$\kappa^*(\bar\pi) = \frac{g_s(\bar\pi, \hat y_d)}{g(\bar\pi, \hat y_d)}.$$

Then for every $\kappa > \kappa^*(\bar\pi)$, $t^*_d < t_s$. Notice that $\kappa > \kappa^*(\bar\pi)$ is a sufficient condition, but not a necessary condition. It may be possible that $t^*_d < t_s$ for some $\kappa \in (1, \kappa^*(\bar\pi)]$.

Next, we show that $\kappa^*(\bar\pi) \in (1, 2)$. First of all, the ratio

$$\frac{g_s(\bar\pi, y)}{g(\bar\pi, y)} = \frac{\bar\pi e^{2y} + (1-\bar\pi)(y+1)^2}{\bar\pi e^{y}(y+1) + (1-\bar\pi)(y+1)^2} > 1, \ \forall\, y > 0, \tag{8.21}$$

because $e^y > 1 + y$.

Next, it follows from (8.17) that $\hat y_d$ is given by

$$(1-\bar\pi)(\hat y_d + 1)^2 = \bar\pi e^{2\hat y_d}(2\hat y_d^2 - 1),$$

and $\hat y_d \in (1/\sqrt{2}, 1)$. Using this and (8.21), we calculate

$$\begin{aligned} \frac{g_s(\bar\pi, \hat y_d)}{g(\bar\pi, \hat y_d)} &= \frac{\bar\pi e^{2\hat y_d} + (1-\bar\pi)(\hat y_d+1)^2}{\bar\pi e^{\hat y_d}(\hat y_d+1) + (1-\bar\pi)(\hat y_d+1)^2} \\ &= \frac{2\hat y_d^2 e^{\hat y_d}}{\hat y_d + 1 + e^{\hat y_d}(2\hat y_d^2 - 1)}. \end{aligned}$$

To establish the fact that $\kappa^*(\bar\pi) < 2$, we need to show that

$$\begin{aligned} 1 &> \frac{\hat y_d^2 e^{\hat y_d}}{\hat y_d + 1 + e^{\hat y_d}(2\hat y_d^2 - 1)} \\ &\Updownarrow \\ \hat y_d^2 e^{\hat y_d} &< \hat y_d + 1 + e^{\hat y_d}(2\hat y_d^2 - 1) \\ &\Updownarrow \\ 0 &< \hat y_d + 1 + e^{\hat y_d}(\hat y_d^2 - 1) \\ &\Updownarrow \\ 0 &< 1 + e^{\hat y_d}(\hat y_d - 1) \\ &\Updownarrow \\ 1 - \hat y_d &< e^{-\hat y_d}, \end{aligned}$$

which holds for any $\hat y_d > 0$. The inequality $t_s(A, \bar\pi) < t_f(A, \pi_s(\bar\pi, t^*_d))$ follows from the fact that the optimal stopping time is increasing in the initial beliefs and $\bar\pi < \pi_s(\bar\pi, t^*_d)$.

It remains to show that if $t^*_d(A, \bar\pi) < t_s(A, \bar\pi)$, then the leader's value in Theorem 4.15 is smaller than the value of the single experimenter. Let player i be the leader, and player j be the follower. We write

$$\begin{aligned} V_s(\bar{\pi}) &= S + C\int_0^{t_s} e^{-rt} p_s(\bar{\pi}, t)\,(A - \lambda_s(\bar{\pi}, t))\,dt, \\ V_i(\bar{\pi}, 0, \hat{q}_i, \hat{q}_j) &= S + C\int_0^{t_d^*} e^{-rt} p_d(\bar{\pi}, t)\,(A - 0.5\kappa\lambda_d(\bar{\pi}, t))\,dt, \end{aligned}$$

By Lemma 8.4, for $\kappa \geq \kappa^*(\bar{\pi})$ and $t < t_d^*$,

$$A - \lambda_s(\bar{\pi}, t) > A - 0.5\kappa\lambda_d(\bar{\pi}, t),$$

and $p_s(\bar{\pi}, t) > p_d(\bar{\pi}, t)$ for all t. Hence $V_s(\bar{\pi}) > V_i(\bar{\pi}, 0, \hat{q}_i, \hat{q}_j)$.